DOCUMENT RESUME

ED 201 151 EC 132 581

AUTHOR Horner, Robert H., Ed.: Bellamy, G. Thomas, Ed.
TITLE Habilitation of the Severely and Profoundly Retarded: Volume III. Reports from the Specialized Training Program.
INSTITUTION Oregon Univ., Eugene. Rehabilitation Research and Training Center in Mental Retardation.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.: Office of Human Development (DHEW), Washington, D.C. Office of Developmental Disabilities.: Social and Rehabilitation Service (DHEW), Washington, D.C.
PUB DATE Aug 80
GRANT 16-P-56817/0-12; 51-P-50270/9-03; G007605455
NOTE 224p.; For previous monographs in this series, see ED 149 528 and ED 161 171.

EDRS PRICE MF01/PC09 Plus Postage.
DESCRIPTORS *Behavior Modification; *Evaluation Methods; Leisure Time; Life Style; Observation; Operant Conditioning; Productivity; *Program Evaluation; Reinforcement; *Severe Mental Retardation; *Vocational Rehabilitation; Work Sample Tests

ABSTRACT

Sixteen papers concerned with vocational habilitation of the severely and profoundly retarded are organized into three parts: issues in habilitation, stimulus control of vocational behavior, and instrumentation for program evaluation. Papers have the following titles and authors: "Structured Employment: Productivity and Productive Capacity" (R. Horner, G. Bellamy); "A Strategy for Programming Vocational Skills for Severely Handicapped Youth" (G. Bellamy, et al.); "Summer Jobs for Vocational Preparation of Moderately and Severely Retarded Adolescents" (J. Clarke, et al.); "Information Processing for Habilitation" (S. Boles, J. Boomer); "Stimulus Control Within Operant Chains: Variables Affecting the Vocational Performance of a Severely Retarded Woman" (R. Horner, G. Bellamy); "The Effect of Reinforcement Schedules on the Intra-Chain Response Patterns of Three Severely Retarded Workers" (R. Horner, G. Bellamy); "Increasing the Productivity of a Profoundly Retarded Sheltered Workshop Employee" (B. Wuerch); "The Effect of Contingent Praise on the Work Rate of a Severely Retarded Adult" (N. Rumble, et al.); "Systematic Correction Procedures in Vocational-Skill Training of Severely Retarded Individuals" (D. Close, et al.); "The Effects of Immediate Correction on the Error Rate of One Severely Retarded Worker" (R. Horner, N. Prill); "The Intra-Chain Response Patterns of Three Severely Retarded Adults, During Extinction" (R. Horner, et al.); "Multiple Perspective Evaluation of an In-Service Program for Vocational Trainers of Severely Retarded Adults" (N. Weissman-Frisch, et al.); "A Telephone Technique for Assessing Living Environments and Life Styles of Retarded Adults" (N. Weissman-Frisch, et al.); "NOLE: The Naturalistic Observation of Severely Retarded Vocational Trainees in Non-Vocational Leisure Time Settings" (J. Boomer); "The Trainee Performance Sample: Improving Vocational Skill Assessment of Severely Retarded Adults" (L. Irvin, et al.); and "Multiple Assessment Evaluation of Programs for Severely Retarded Adults" (L. Irvin, et al.) (DB)

U S DEPARTMENT OF HEALTH.
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED201151

HABILITATION OF THE SEVERELY AND PROFOUNDLY RETARDED:

Volume III

Reports From The Specialized Training Program

Edited by

Robert H. Horner
G. Thomas Bellamy

Joint Publication

by

Specialized Training Program
Monograph No. 3
and
Rehabilitation Research and Training Center in Mental Retardation
Monograph No. 12

Publication of this Monograph was supported by the Rehabilitation Research and Training Center in Mental Retardation and the Center on Human Development

College of Education
Center on Human Development
University of Oregon
Eugene, Oregon 97403

August 1980

EC132581

This Monograph was prepared without copyright to facilitate
rapid professional dissemination

Specialized Training Program
and
Rehabilitation Research and Training Center in Mental Retardation
University of Oregon

The Social and Rehabilitation Service and the Office of Education
reserve a royalty-free, non-exclusive, and irrevocable license
to reproduce, publish, or otherwise use, and to authorize
others to use, all copyrightable or copyrighted
material resulting from this grant
supported research.

The University of Oregon, a member of the Oregon State System of Higher Education, probhibits discrimination based on race, color, religion, sex, age, handicap or national origin. This policy implements various federal and state laws, Executive Orders, Title IX and its regulations, and applies to employment, admission, education and facilities. Direct related inquiries to Norma C. McFadden, Director, Affirmative Action Office, 472 Oregon Hall, University of Oregon, (503) 686-3123.

This publication was supported in part by grants to the Specialized Training Program from the Developmental Disabilities Office, Region X (Grant Number 51-P-50270/9-03) and from the Bureau of Education for the Handicapped, Research Grant Number G00-760-5455, and by Rehabilitation Research and Training Center Grant 16-P-56817/0-12 from Social and Rehabilitation Service, Department of Health, Education and Welfare, Washington, D. C. 20201.

Table Of Contents

List of Contributors -- v

Preface -- vi

PART I: ISSUES IN HABILITATION

Structured Employment: Productivity and Productive Capacity -- 3
R. H. Horner and G. T. Bellamy

A Strategy for Programming Vocational Skills for Severely Handicapped Youth---------------------------------- 19
G. T. Bellamy, D. J. Wilson, E. Adler, and J. Y. Clarke

Summer Jobs for Vocational Preparation of Moderately and Severely Retarded Adolescents------------------------ 35
J. Y. Clarke, L. Greenwood, D. Abramowitz, and G. T. Bellamy

Information Processing for Habilitation-------------------- 49
S. M. Boles and J. D. Boomer

PART II: STIMULUS CONTROL OF VOCATIONAL BEHAVIOR

Stimulus Control Within Operant Chains: Variables Affecting the Vocational Performance of a Severely Retarded Woman-- 61
R. H. Horner and G. T. Bellamy

The Effect of Reinforcement Schedules on the Intra-Chain Response Patterns of Three Severely Retarded Workers----- 83
R. H. Horner and G. T. Bellamy

Increasing the Productivity of a Profoundly Retarded Sheltered Workshop Employee------------------------------ 99
B. B. Wuerch

The Effect of Contingent Praise on the Work Rate of a Severely Retarded Adult------------------------------- 109
N. Rumble, D. P. Inman, and R. H. Horner

Systematic Correction Procedures in Vocational-Skill Training of Severely Retarded Individuals----------- 121
D. W. Close, L. K. Irvin, H. J. Prehm and V. E. Taylor

The Effects of Immediate Correction on the Error Rate of One Severely Retarded Worker--------------------- 131
R. H. Horner and N. M. Prill

The Intra-Chain Response Patterns of Three Severely Retarded Adults During Extinction--------------- 139
R. H. Horner, B. B. Wuerch, and J. D. Boomer

PART III: INSTRUMENTATION FOR PROGRAM EVALUATION

Multiple Perspective Evaluation of an In-Service Program for Vocational Trainers of Severely Retarded Adults-- 153
N. Weissman-Frisch, F. A. Crowell, and D. P. Inman

A Telephone Technique for Assessing Living Environments and Life Styles of Retarded Adults---------- 169
N. Weissman-Frisch, F. A. Crowell, G. T. Bellamy, and D. Bostwick

NOLE: The Naturalistic Observation of Severely Retarded Vocational Trainees in Non-Vocational Leisure Time Settings------------------------------------ 185
F. A. Crowell, N. Weissman-Frisch, D. Lamson, and J. D. Boomer

The Trainee Performance Sample: Improving Vocational Skill Assessment of Severely Retarded Adults-- 199
L. K. Irvin, R. Gersten, G. T. Bellamy, V. Taylor and D. W. Close

Multiple Assessment Evaluation of Programs for Severely Retarded Adults-------------------------------- 217
L. K. Irvin, F. A. Crowell, and G. T. Bellamy

List Of Contributors

David B. Abramowitz is Special Needs Coordinator with the Lane County Department of Employment and Training in Eugene, Oregon.

Ellen Adler, M.S., is a Specialist, TMR Program with the Lane Education Service District, Eugene, Oregon.

G. Thomas Bellamy, Ph.D., is Director of the Specialized Training Program, and an Associate Professor of Special Education at the University of Oregon.

Shawn M. Boles, Ph.D., is the Evaluation Coordinator of the Specialized Training Program, University of Oregon.

J. Daniel Boomer, M. S., is a Research Assistant with the Specialized Training Program, University of Oregon.

David Bostwick, Ph.D., was a Research Assistant with the Rehabilitation Research and Training Center at the University of Oregon.

James Y. Clarke is a Research Assistant with the Specialized Training Program, University of Oregon.

Daniel W. Close, Ph.D., is an Assistant Professor in Special Education and a Research Associate with the Rehabilitation Research and Training Center at the University of Oregon.

Fred A. Crowell, Ph.D., is a Research Associate with the University of Oregon.

Lynn M. Greenwood, B.A., is a masters student in Special Education at the University of Oregon.

Robert H. Horner, Ph.D., is Associate Director of the Specialized Training Program and an Assistant Professor of Special Education at the University of Oregon.

Dean P. Inman, Ph.D., is a Research Asscoiate with the Center on Human Development at the University of Oregon.

Larry K. Irvin, Ph.D., is a Research Associate with the Rehabilitation Research and Training Center and an Assistant Professor of Special Education with the University of Oregon.

Deborah Lamson, M.A., was a Research Assistant with the Specialized Training Program, University of Oregon.

Herbert J. Prehm, Ph.D., is a Professor of Special Education at Arizona State University.

Nancy M. Prill, M.S., is a teacher of Special Education.

Newt Rumble, M.S., is Vocational Evaluator for United Cerebral Palsy in Seattle, Washington.

Valerie E. Taylor, M.S., is a Research Assistant with the Rehabilitation Research and Training Center at the University of Oregon.

Nancy Weissman-Frisch, M.S.S.W., M.S., is Inservice Coordinator with the State of Oregon's Division of Mental Retardation and Developmental Disabilities.

Darla J. Wilson is a Research Assistant with the Specialized Training Program, University of Oregon.

Bonnie Biel Wuerch, M.S., formally a Research Assistant with the Specialized Training Program, is currently a Research Assistant with the University of Hawaii at Manoa.

Preface

This is the third in a series of monographs compiled by the University of Oregon's Specialized Training Program to facilitate broad dissemination of research and service reports. The purpose of the Specialized Training Program is to increase vocational opportunities for severely and profoundly retarded adults in the Pacific Northwest. Strategies for accomplishing this include: (a) development of a comprehensive model for workshop operation, (b) management of a demonstration workshop at the University of Oregon, (c) extensive assistance to facilities in the Northwest who adopt the model, (d) inservice training, and (d) research.

Papers in this and the previous two volumes document a chronology of practical and conceptual problems we have encountered or anticipated in providing vocational services to severely handicapped adults. These problems have guided our efforts to develop solutions through research and demonstration.

The papers in this monograph reflect our interest in continuing to address multiple service delivery issues. The first section of the monograph includes several discussions of current issues affecting vocational services to severely handicapped individuals. Significant policy decisions will be made at all governmental levels during the next few years as severely handicapped people increasingly expect community-based services. The interplay of Congressional and court actions, research results, current service models, funding priorities, and changing values of professionals and advocates provides the clay from which these decisions will be modeled. Chapters in section one continue our grappling with these issues.

The papers in Section 2 describe ongoing behavioral research to identify procedures for vocational training and supervision. In a recent textbook, (Bellamy, Horner, & Inman, 1979) we suggested that a technology of vocational habilitation is now available for severely handicapped people. Despite the availability of extensive procedures, significant service problems remain. Particular interest during the past year has focused on identification of variables that affect task performance in work settings. The papers presented in Section 2 describe detailed analyses of stimulus control within vocational settings.

The third section of the monograph documents our continuing interest in the development of assessment instruments that can assist in evaluating the effects of vocational services. Increasing emphasis on the development of dispersed community service makes it difficult to evaluate effectiveness of service efforts. Practical measurement tools are needed to assist in evaluation of both individuals and programs.

This monograph and the work reported in it were made possible by the support of several agencies and individuals. The Specialized Training Program was supported in the research by the Department of Health, Education, and Welfare Developmental Disabilities Office, Region X and the Bureau of Education for the Handicapped. Preparation of the text was greatly assisted by the clerical effort and skills of Theresa Bush and Dorothy Sweet and the graphics work by Arden Munkres and Vanessa Tsang.

Robert H. Horner
G. Thomas Bellamy
May, 1980

PART I

Issues in Habilitation

Structured Employment: Productivity and Productive Capacity

Robert H. Horner and G. Thomas Bellamy

University of Oregon

This research was supported by the DHEW Developmental Disabilities Office, Region X, and the Bureau of Education for the Handicapped, Research Branch.

This chapter originally appeared in G. Bellamy, G. O'Connor, and O. Karan, Vocational rehabilitation of severely handicapped persons, 1979, University Park Press. Reprinted with permission.

Structured Employment : Productivity and Productive Capacity

This paper focuses simultaneously on two topics, work productivity and severely retarded adults. To date, the two have had little in common, except that the presence of one implied the absence of the other. The purpose of this paper is to explore this negative relationship and to suggest a strategy that may be useful in assisting severely handicapped adults to work productively.

Productivity and Productive Capacity

Production involves the use of labor, in conjunction with capital and/or raw materials, to create exchangeable economic value. Productivity indexes the rate with which such value is created. Productivity can be measured with respect to a nation, a corporation, a group or an individual. In this chapter we are interested in the productivity of individuals.

Within the existing social and economic structure productivity plays a critical role in an individual's quality of life. As the "mother of wages" (George, 1881) productivity has much to do with access to work opportunities, to economic independence and to the necessities and amenities of modern living. While there clearly is not a one-to-one relationship between level of productivity and level of remuneration, a minimal level of productivity is typically prerequisite for access to, and maintenance of, employment.

The importance of individual productivity is also apparent in society as a whole. We are currently challenged by our political leaders, for example, to produce our way out of an economic slump. Similarly, one variable consistently pointed to as a factor in our ability to regain financial stability in the world currency market is the level of individual worker productivity.

"Productive capacity" refers to an individual's or organizations' predicted level of productivity at any point in time. While productivity refers to how much was produced, productive capacity refers to an estimate of how much will be produced. The productive capacity of an individual is affected by many factors, including: (a) the type of task being performed, (b) the existing skill level of the worker, (c) the ability of the worker to learn requisite skills not in his or her repertoire, (d) the setting conditions under which training and production take place, and (e) the level of capitalization and automation of the production process. The productive capacity of employees within a work force is important information for managers. To predict work output and resource needs (i.e., how much raw material and/or capital will be required for a job), a manager must rely on some index of each worker's productive capacity. Employers only hire workers they believe have the capacity to be productive. As such, individuals perceived as having "limited" or "low" productive capacity face restricted opportunities for employment.

A point worth stressing in this definition of productivity is that an individual's productive capacity is not static. It changes over time in response to many different variables. One important change in individual productive capacity results as a function of training and experience. The productive capacity of an individual employed as a mechanic, for instance, is relatively low until s/he has received training in the complex, interrelated systems that propel vehicles. Following training, the individual's level of productivity typically is much higher. His or her productive capacity changes as a function of training. Similarly, the same individual may become increasingly competent with practice (i.e., over time the same job is done faster). Here again productive capacity changes. The worker's level of productivity improves as a function of experience or practice.

Because productivity changes as a result of the job demands, training and experience, it does not seem useful to consider productivity as a stable personality trait. An individual does not "possess" low productivity any more than s/he "possesses" a third grade reading ability or a particular level of physical conditioning. While each may be descriptive of current performance, each may also be affected by access to opportunities for learning and practice. Therefore, such characteristics are not necessarily predictive of long term performance.

Severely Handicapped Adults

One group of individuals who typically have been perceived as "possessing" limited productive capacity are severely handicapped adults. For the purposes of this chapter we will define severely handicapped adults as those individuals whose level of retardation or multiple handicaps can be characterized by some or all of the following: (a) IQ scores below 40, (b) pronounced skill deficits in personal, social and vocational domains, (c) a history of exclusion from training opportunities that might lead to vocational competence, and (d) a need for substantial training before behavioral changes are achieved (Gold, 1975). The relationship between these characteristics and direct intervention strategies has been presented in detail by Bellamy, Horner, and Inman (1979).

When viewed from an economic perspective an additional characteristic of severely handicapped adults is apparent: they generally function only as consumers in our society. Most non-handicapped adults have roles as both consumers and producers. However, severely handicapped adults have not been perceived as possessing the potential (i.e., productive capacity) to fulfill the role of a producer. Few have access to competitive employment options. Those in sheltered employment earn on the average, only trivial wages (Whitehead, 1978), and many are excluded from vocational opportunities altogether (Lynch & Graber, 1977). This low productivity is reflected in restricted personal financial options and dependence on income transfer programs for basic necessities.

Productivity and Severely Handicapped Adults

With recognition of the civil rights and learning ability of handicapped adults has come a feeling that productive work is both desirable and possible for many severely handicapped individuals. Productivity is desirable because it (a) increases the economic independence of handicapped adults; (b) allows them to assume a more normative adult role; (c) decreases their dependence on public support; and (d) may increase the likelihood of acceptance by non-handicapped peers (Gold, 1975). The desirability of productive work as a goal for severely handicapped adults has been recognized by advocates of normalization (Wolfensberger, 1972; Olshanski, 1972). This may stem in part from increasing awareness of the role work plays as a basic element of normal lifestyles (Turkel, 1972).

In addition to being desirable, productivity is possible for severely handicapped adults. There exists a rapidly expanding professional literature which documents the learning of difficult, lengthy, and intricate vocational skills by severely handicapped individuals. During the last few years research has left little doubt that severely handicapped adults can become vocationally competent persons (Bellamy, 1976a; Bellamy, Horner & Inman, 1977; Gold, 1976; Jacobs, 1976, 1978; Karan, Wehman, Renzaglia & Schutz, 1976). Thus, the literature documenting low productivity by severely handicapped adults (U.S. Department of Labor, 1977) should not be viewed as a reliable index of the productive capacity of these individuals.

The Discrepancy Between Productivity and Productive Capacity

At present we are faced with encouraging research results which have had little impact on the day-to-day vocational options of severely handicapped individuals. Although repeated demonstrations of learning competence exist, severely handicapped adults continue either to earn insignificant wages or to face complete exclusion from vocational training and employment options. There are probably many reasons for this discrepancy. For example, much has been said recently about low professional and public expectations (Gold, 1975); lack of adequate work in sheltered workshops (Whitehead, 1979); failure of workshops to include many severely handicapped individuals (Greenleigh Associates, 1975) disincentives in the social service system for high earning levels (Kosmo & Pritchard, undated; Pomerantz & Marholin, 1977); and the expense required for habilitation of severely handicapped clients (Conley, 1973; Levithan & Taggart, 1977). Two additional factors, that have received less attention, seem inexorably linked to the discrepancy between the expected and observed vocational abilities of severely handicapped workers. The first is the lack of research efforts to document the long-range productivity and earnings of severely handicapped individuals following training. The second is recognition that organizational and administrative barriers limit the access of severely handicapped individuals to current employment options.

A Need for Earning Demonstrations. While acquisition and short-term maintenance of a skill may be an acceptable level of "success". Skill acquisition is important and functional only if it is related to consistent, productive behavior on remunerative tasks. The current literature has not demonstrated such a relationship. Vocational research with severely handicapped adults has documented only the learning potential of these individuals, not their earning potential The vocational skills learned have been impressive, but the opportunity to use these skills in extended remunerative employment has not been documented adequately. Future research efforts are needed that focus on the long-range earning capabilities of severely handicapped adults.

Access to Employment Options. A repeated problem faced by advocates for the severely handicapped is that the vocational skills reported in the research literature are only a subset of those presently required for significant individual productivity. Deficiencies in skills that are indirectly related to productivity often prevent a worker from having access to lucrative work situations. It is consistent with the U. S. Department of Labor's (1977) data on low annual wages of retarded adults in sheltered workshops to conclude that competitive employment offers the only current opportunity for adequately paid work. Clearly, the acquisition of task-specific skills is insufficient for competitive work, where changing job demands, highly verbal supervision procedures, and absence of support outside the work setting may be expected. With the onset of individual productivity after competitive job placement, most rehabilitation and social service supports are typically withdrawn, and the individual is expected to work and live rather independently. This model of "train, place and forget" may be appropriate for many mildly handicapped individuals whose entry or reentry into the labor force requires only a one-time, short-term intervention (although Edgerton & Bercovici's (1976) data suggest that frequent job changes, productivity interruptions and resulting lifestyle disruptions are characteristic of many of their mildly retarded adult subjects).

Such a "train, place and forget" tactic is clearly inappropriate for severely handicapped individuals. This is true for at least two reasons: (a) "Training" is unlikely to result in cost effective placement of severely handicapped workers as long as the only available employment opportunities require independence on skills both directly and indirectly related to work performance; and (b) "Forgetting" an individual after placement fails to provide for ongoing personal and work-related support that may be necessary to maintain the independence and productivity of a severely handicapped worker.

Existing adult services reflect the assumption that only those individuals who can perform skills indirectly related to work (i.e., independent grooming, dressing, communicating, traveling and eating) can be expected to perform those direct vocational skills associated with remunerative employment options. If this assumption is inaccurate, and considerable research suggests that it is, then a large number of potentially productive individuals are being unnecessarily denied access to meaningful employment. It is our contention that the current restrictive nature of employment options, and the service system which has developed to coordinate with them, significantly reduce the productivity of severely handicapped adults. Therefore, the structure of employment options, not just the disabilities of individual workers, can be viewed as a causal factor in low productivity.

Few employment or social service structures provide long-term support for individuals in competitive employment. It is therefore hardly surprising that independent living and personal-social skills often are treated by rehabilitation professionals as prerequisites for job training or other vocational services. It is our view that such "prerequisites" reflect only the present restricted range of employment opportunities, and are not related to any necessary skill sequence or learning hierarchy in individual development. If structured employment opportunities were provided so that high productivity levels could be attained while support and training on indirect skills continued, vocational rehabilitation agencies probably would not find it necessary to exclude so many severely handicapped people.

Structured Employment

If the low productivity levels of severely handicapped adults are attributed at least in part to the structure of employment opportunities, efforts to improve productivity should target changes in long term employment, not just improvements in briefer rehabilitation or training efforts. For this reason we suggest that particular attention be given to planning employment opportunities in which severely handicapped individuals can be highly productive without needing first to demonstrate mastery of social and independent living skills that are only indirectly related to vocational competence.

This has been the objective of many extended employment programs in sheltered workshops (Olshanski, 1977), but Whitehead's (1979) report that the average annual wage was $417 for retarded employees of sheltered workshops suggests that continued analysis of extended employment mechanisms is needed. Our concept of structured employment is based on (a) the assumption that low individual productivity and earnings often reflect availability of work and adequacy of supervision, rather than disability of the worker; and (b) that with carefully designed work environments, normal productivity and normal wages are possible for many severely handicapped individuals who are now constrained to participation in "prevocational" programs.

Structured employment is a type of extended employment that emphasizes high levels of individual productivity with ongoing training and support services not typically provided in industry. As we define it, structured employment involves vocational opportunities that (a) focus on severely handicapped individuals; (b) provide a long-term employment option; (c) emphasize high productivity with commensurate wages; (d) provide ongoing employment support; and (e) reflect administrative breadth.

Focus on severely handicapped individuals

It is our assumption that structured employment options will be most appropriate for severely and profoundly handicapped individuals. Research and service efforts are demonstrating that most mildly and moderately handicapped people can learn the direct and indirect skills needed for competitive placement (e.g., Sowers, Thompson, and Connis, 1979). We would also suggest that the currently available, time-limited rehabilitation options are designed primarily to meet the needs of mildly handicapped persons. The structured employment option is designed for those lower functioning individuals who are typically perceived as having "inconsequential productive capacity" (Federal Register, 1974) in the present rehabilitation structure.

Long-term employment

The objective of structured employment is to allow severely handicapped workers to use the vocational skills they have learned. Individual placement in industry is a lesser concern than is competent, adequately remunerated performance within the structured employment setting. Thus, structured employment provides a heretofore missing link within the current continuum of services. Handicapped individuals who both have skills directly related to production tasks and skills indirectly related to production (grooming, hygiene, communication, self-help, etc.) can be placed in competitive employment settings. These settings afford the least restrictive alternative for involvement in the "world of work" for these individuals. Handicapped workers who have indirect but not direct skills enter sheltered environments designed to provide for the training of direct skills. Following training the workers are placed in a competitive setting. However, individuals who are able to perform, or learn to perform direct vocational skills, but lack many indirect skills are frequently excluded from vocational opportunities. A structured employment option would allow these individuals to actualize their productive abilities while receiving additional training (during non-working hours) focused on their indirect skill deficits. Should a worker achieve proficiency in both indirect and direct skills s/he would become a candidate for competitive employment. In the interim the person would have the opportunity to use the vocational competence s/he has. Therefore, providing the opportunity for a handicapped worker to work over long time periods with specialized support simply reflects an administrative recognition that (a) performance of complex social, communication and self-help skills is not a necessary prerequisite for access to normally productive employment, and (b) rapid acquisition of these indirect skills need not be contingently related to maintaining that access. The long-term employment option provided by structured employment could decrease the job discrimination against competent workers who have skill deficits that are only indirectly related to on-the-job performance.

High productivity and commensurate wages

Structured employment is first and foremost an employment option. The "service" being provided is support or subsidy for an administrative structure that allows access to employment for lower functioning individuals. Structured employment does not fall within the framework of the traditional social service system. The primary indices of the program effectiveness are the productivity level of individual workers and the wages that this productivity leads to. As such, structured employment options should allow workers access to tasks with non-trivial remunerative potential. Following Crosson's (1969) oft-cited recommendation, work should be chosen based on its economic value not the ease with which it can be trained.

A central assumption of structured employment is that one of the greatest services we can provide a severely handicapped adult is access to meaningful work. It is suggested here that access to a viable social role in a community and the "attitude" of caretakers, parents, and neighbors toward a severely handicapped person will be affected by that person's ability to demonstrate competence in activities valued by community members. Given the status that work holds for adults in our society, a major benefit might be expected if severely retarded individuals were given the opportunity to perform successfully in a socially acceptable, adult-appropriate work role. Such demonstrations of competence will not occur if a worker's rate of production is trivially low or if the wages generated by high productivity are inconsequential.

Ongoing employment support

The central success measure in structured employment is the average income (level of productivity) of workers over time. Providing a work setting in which severely handicapped workers can maintain high productivity levels for extended periods requires support for two activities that may not be economically feasible for competitive industries.

The first is access to a production setting with specialized supervision. The second is the opportunity for further training after placement in the production setting. Production supervision techniques are a critical variable affecting the productivity of severely handicapped workers (Bellamy, 1976b; Martin & Pallotta, 1978). Those few studies which have demonstrated prolonged, high-rate production by severely handicapped workers have typically used specialized production supervision procedures (Bellamy, Inman & Yeates, 1978). As such, structured employment would be characterized by ongoing specialized supervision as the trained worker entered production. We consider this extended supervision to be one of the most important characteristics of a structured employment option, and will consider it in greater depth later in this chapter.

The second element of ongoing service support is continued access to training after a worker enters production. Workers in a production environment may, over time, (a) begin to perform unacceptably on the task or (b) complete the contract they were trained for and have no other work that requires their specific skills. In either case additional training may be critical to the worker's continued productivity. The first type of problem has been described by Bellamy. Inman and Horner (1978) and Bellamy, Horner and Inman (1979). These authors indicate that systematic retraining is one option that has been successful when a worker's production behavior gradually becomes unacceptable.

With the second type of problem, completion of a contract, a worker may be left with vocational skills that often become obsolete overnight. Here again the most obvious solution is additional training on new contracts. Given the specialized training required by most severely handicapped people it is reasonable to expect that any viable employment environment will need to have the ongoing capability of retraining workers.

We see ongoing service support as a major factor affecting the long term success of severely handicapped workers. Such support is based on the recognition that both the skill level of a worker and his or her access to work may vary over time. Within structured employment a worker would receive initial training on a task, and then perform that task as long as it was economically worthwhile. With the completion of the task, training would begin on another job. It would, of course, also be possible for a worker to receive training on other tasks while continuing to spend most of his or her workday in production. As one job ended, the worker would then move rapidly into production on another.

To provide specialized supervision and periodic retraining, an organization offering structured employment may need to rely on public support for supervision costs not normally encountered in private industry. Although few mechanisms now exist for providing this support, there is a growing interest in extended support options in vocational facilities (Committee for Development of Direct Long Term Funding for Workshops and Work Activity Centers, 1977).

Administrative breadth

Many factors affect worker productivity. Those mentioned above are but a few that we feel are particularly important for severely handicapped workers. It is recognized, however, that structured employment, like any work option, should incorporate the entire spectrum of administrative and business needs facing a workshop. Whitehead (1979) emphasized that workshops should attend simultaneously to a variety of business and service considerations. These include (a) sales; (b) engineering; (c) finance; (d) management; (e) training; and (f) supervision. Within a highly competitive business climate it is unlikely that a workshop (employing either handicapped or non-handicapped workers) will be able to demonstrate prolonged, high levels of worker productivity without attention to all of the above. As such, structured employment options will need to provide the administrative breadth needed to cover each of these areas.

While this would seem a mammoth task, it is important to recognize that the sales, engineering, finance, management, training and supervision problems being faced by workshops are similar to those dealt with for decades by the business community. In fact the above list of adminstrative foci reads much like the corporate organizational chart of any major business. As such, models for dealing with these issues are quite likely to be found in successful businesses. The appropriateness of these models gains credence when one recognizes that one objective of a business is likely to be the same as that of a structured employment option: to optimize the opportunity for workers to be productive.

We join Whitehead (1979) in suggesting that workshop models for sales, engineering, finance and management can best be found in the experienced roots of the business community. Each of these areas deal with staff-to-public or staff-to-staff interactions. The staff roles and objectives which are involved when a workshop performs sales, engineering, finance and management functions are substantially the same roles and objectives which are involved in any successful business venture.

This logic does not, however, extend to the areas of training and supervision. The training and supervision strategies employed by competitive business are designed specifically for non-handicapped individuals. The learning options that these strategies afford are exactly those in which severely handicapped individuals have failed. As such, we suggest that strategies for training and supervising handicapped workers will best come from a more detailed analysis of behavior in the work setting. A major task before the field of habilitation, therefore, is to develop a technology that both focuses on training and supervision strategies suitable for severely handicapped workers, and integrates these strategies with ongoing business expertise.

We present structured employment as an alternative vocational option for severely handicapped individuals that emphasizes high productivity with specialized support. Structured employment reflects the assumption that severely handicapped adults have both the ability and the right to work. The objective for service providers is to identify and develop an environment that will facilitate significant productivity. Each of the characteristics listed above is designed to serve this objective. Long-term employment, a focus on severely handicapped workers, high productivity, commensurate pay, ongoing support, and administrative breadth all affect a severely handicapped worker's ability to function successfully in an employment setting.

It is important to note, however, that these characteristics do not limit employment of severely handicapped individuals to a sheltered workshop. They simply outline the type of support needed for productive employment of severely handicapped workers. This support could take place in extended employment programs in workshops, within an enclave established in an industry (DuRand & Neufeldt, 1975), in a worker-owned cooperative, in a private enterprise, or in any of a variety of administrative structures that afford ongoing support.

A Production Supervision Technology for Structured Employment

Structured employment is unique in its combined focus on (a) long-term employment, (b) high productivity levels, and (c) severely handicapped people. For structured employment to provide a viable work option, a production supervision technology is needed, which provides the ongoing support that allows severely handicapped employees to be highly productive. The remainder of this chapter will outline characteristics of such a supervision technology and describe one recent study which has contributed to the area of production supervision.

Of course, our focus on production supervision as a key element in structured employment reflects an assumption that severely handicapped individuals can learn to perform requisite skills. Research focusing on training such skills has now demonstrated that task acquisition is not only possible but probable (Gold, 1976). The present volume presents a variety of successful approaches to vocational training with the severely handicapped, and Bellamy, Horner, and Inman (1979) have recently provided a detailed description of a technology of training that has been effective with many severely handicapped people. While further refinement and development of training techniques is needed, the major training issue facing the habilitation field today is not whether procedures will be effective, but whether training will lead to remunerative employment options.

The area of production supervision has been addressed by many researchers. Reviews of this research have been provided by Bellamy (1976b), Gold (1973) and Martin and Pallotta (1978). These reviews indicate that (a) production studies have typically employed mildly and moderately handicapped workers and subjects with only a few focusing specifically on the severely handicapped, (b) some techniques exist for some problems, but that (c) critical variables related to long-term maintenance and demonstrated economic output have not been addressed.

As Bellamy, Horner and Inman (1979) indicate, many production supervision "techniques" are available. What is lacking is a synthesis of these techniques into a viable technology. Techniques within this context refer to the systematic manipulation of antecedent and/or consequent stimuli to alter the probability of targeted responses (i.e., task performance). Production supervision techniques have focused on the kind, amount and frequency of supervisor assistance provided prior to worker responding, and the type and schedule of supervisor provided consequation delivered once the worker does respond. Any systematic (i.e., rule governed) procedure for manipulating these antecedent and consequent variables is a technique.

These techniques will form the foundation of a production supervision technology. At present we see several needs if this technology is to reach fruition: (a) the development and acceptance of additional techniques, many of which are available from the fields of psychology and education; (b) continued refinement and implementation of existing techniques; (c) a commitment to systematic measurement of any production technique; (d) evaluation of a technique based on the data provided by such measurement; (e) a system for dealing with those instances in which program techniques fail to produce desired results (Baer, 1977); and (f) an emphasis on the details of day-to-day production supervision (Bellamy, Horner & Inman, 1979).

Within the context of structured employment a production supervision technology will be evaluated on its effectiveness in allowing severely handicapped workers to be productive. Useful production supervision techniques will be those which are functionally related to improved worker productivity. Specialized supervision is a key element of structured employment for two reasons: (a) the expected performance (i.e., productive capacity) of severely handicapped workers is low in production environments employing conventional supervision techniques; and (b) non-trivial vocational options are unlikely to be available to severely handicapped adults unless specialized supervision techniques can be combined to form a functional supervision technology. This emphasis on specialized supervision in structured employment will be critical if learning demonstrations are to be transformed into long-term earning demonstrations.

Toward a Production Supervision Technology: A Research Example

The complexity and interactive effects of production setting variables on a worker's performance rate has been recently demonstrated by Horner and Bellamy (1978). These authors monitored one severely retarded worker's behavior over an extended time period. The worker was performing a small parts assembly task in the Specialized Training Program production environment (cf. Bellamy, Inman and Horner, 1978). The worker in question was exhibiting two problems. First, her rate of production was unacceptably low; and second, she engaged in a variety of irrevelant behaviors as she performed the task (i.e., she rocked parts back and forth, stroked her face, sat motionless for long time periods and performed exaggerated shrugging responses).

Both of these performance patterns, slow working and a high rate of behaviors irrelevant to work completion, are common in workshops serving handicapped individuals. The research strategy was to identify those stimuli controlling irrelevant responding, use the information to decrease irrelevant response rates, and determine if this reduction affected the worker's productivity level. The first intervention following baseline was to increase the level of reinforcement the worker received for work completion.

Results indicate that the rate of three of the four irrelevant responses substantially reduced when the worker was simply provided with a more powerful reinforcer. That is, when reinforcement for work was increased the rate of three of the irrelevant responses decreased. The worker spent more time on-task and her rate of production improved. Unfortunately, the fourth and most problematic irrelevant response did not decrease. In fact it increased in rate. Careful observation indicated that the worker not only performed this irrelevant response more often, but performed it at very specific points, or steps in the task. It was as if the worker added several irrelevant steps to the task, and was treating these irrelevant steps as critical to task completion. From an operant perspective, the data indicated that this high rate irrelevant response was under control of specific stimuli inherent to the task.

To remedy this situation one supervisor retrained the worker over a two day period. During retraining the irrelevant behavior was interrupted by the supervisor. Correct performance at those steps in the task where irrelevant responding had been common was rewarded. Within two days, the worker was performing the task with virtually no irrelevant responses and her production rate had improved over 400%.

A final problem continued to exist, however. The worker only performed well in the presence of the specific supervisor who retrained her. If any other supervisors were working on the production floor the rate of irrelevant responses increased dramatically and productivity plummeted. Within a few days after retraining it became clear that even though no supervisors were consequating the worker's irrelevant responding these responses were only kept at a minimum if the particular supervisor who did retraining was present. This dilemma was easily dealt with by having all supervisors interrupt the worker's irrelevant behavior. Within a day the worker stopped performing the irrelevant response altogether, and has not reinitiated it over the six months since the program was terminated.

We have described this particular study in some detail because it exemplifies the complexity of a "production environment," the adaptive sophistication of a severely retarded worker, and the type of production technology that we see as imperative if severely handicapped individuals are to achieve substantial production rates.

Environmental complexity is apparent in the fact that many things within the production setting affected worker performance. The type and amount of reinforcement affected some irrelevant responses, but a particular retraining effort was required to deal with another. The worker learned very quickly what cues set the occasion for fast working. Once those cues, and a consistent pattern of reinforcement were established the worker's higher production rate maintained over a substantial time. No one manipulation would have been sufficient to produce this effect. Rather, a combination of interventions was required.

The adaptive sophistication of the worker is illustrated by her rapid shift in performance following interventions and the subtle discrimination she made among supervisors, consequences and irrelevant responses. It is of critical importance for individuals responsible for production environments to recognize that worker behavior will largely be an adaptation to the "natural" contingencies in the production setting. If off-task behavior is rewarded (via peer attention, staff attention, access to self stimulation, etc.) more than on-task behavior, the worker will adapt by performing high rates of off-task behavior. The worker in the Horner and Bellamy (1978) study responded quickly when changes were made in her environment. It took her only a few trials to respond to the change in reinforcement schedule. Within ten minutes her behavior during retraining shifted. Perhaps most impressive was her discrimination among individual staff members. Though labeled severely retarded and functioning at Adaptive Behavior Level IV, the worker was very capable of tracking exactly who was supervising and when supervisors changed.

Finally, the study emphasized application of production techniques which are firmly based on a functional, as opposed to procedural, approach to habilitation (cf. Bellamy, Horner & Inman, 1979). Each intervention in the study (i.e., the improved reinforcement, retraining, staff consequation) was a response to measured worker behavior. Each intervention was anticipated to be in-and-of-itself effective in remediating the targeted problems. Each intervention was evaluated specifically on the effect it had on worker behavior. As such, the study represents an example of using measurement and attention to detail which must occur if terms such as "habilitation" and "vocational competence" are to become synonomous with substantial vocational options for severely handicapped citizens.

Summary

Current research with severely handicapped workers indicates that these individuals can learn complex vocational tasks. If this research is to have significant impact on the vocational options available to severely handicapped people, attention must also be directed at environmental characteristics that will promote long-term performance of learned skills. Severely handicapped adults can be productive, but are often excluded from most work environments because of skill deficits only tangentially related to work. Structured employment is presented as a framework that will facilitate and support that potential productivity. Characterized by long-term employment, a focus on severely handicapped individuals, high levels of productivity with commensurate wages, ongoing service support and administrative breadth, structured employment is seen as appropriate for a wide range of employment settings. Central to the continued refinement of structured employment options is emphasis on a technology of production supervision in which techniques are designed for individual workers and evaluated by continuous monitoring of worker productivity.

References

Baer, D. M. Prepared remarks at the meeting of the American Association for the Advancement of Behavior Therapy, Atlanta, December, 1977.

Bellamy, G. T. Habilitation of the severely and profoundly retarded: A review of research on work productivity. In T. Bellamy (Ed.), *Habilitation of severely and profoundly retarded adults, Vol. 1.* Eugene, Oregon : University of Oregon Center on Human Development, 1976b.

Bellamy, G. T. (Ed.), *Habilitation of severely and profoundly retarded adults, Vol. 1.* Eugene, Oregon: University of Oregon Center on Human Development, 1976a.

Bellamy, G. T., Horner, R. H. and Inman D. P. *Vocational habilitation of severely retarded adults: A direct service technology.* Baltimore: University Park Press, 1979.

Bellamy, G. T., Horner, R. H. and Inman, D. P. (Eds.) *Habilitation of severely and profoundly retarded adults: Vol. II.* Eugene, Oregon: University of Oregon Center on Human Development, 1977.

Bellamy, G. T., Inman, D. P. and Horner, R. H. Design of vocational habilitation services for the severely retarded: The Specialized Training Program model. In G. Hamerlynch (Ed.), *Applied behavior analysis techniques for the developmentally disabled.* New York: Brunner/Mazel, 1978.

Bellamy, G. T., Inman, D. P. and Yeates, J. Evaluation of a procedure for production management with the severely retarded. *Mental Retardation*, 1978, *16* (4), 317-319.

Committee for Development of Direct Long Term Funding for Workshops and Work Activity Centers. *The case for federal support for extended sheltered workshop services for community-based, severely disabled, substantially handicapped adults*, 1977.

Conley, R. *The economics of mental retardation.* Baltimore: Johns Hopkins Press, 1973.

Crosson, J. A technique for programming sheltered workshop environments for training severely retarded workers. *American Journal of Mental Deficiency*, 1969, *73*, 814-818.

DuRand, J. and Neufeldt, A. H. Comprehensive vocational service systems. *Monograph 4*, National Institute on Mental Retardation, 1975.

Edgerton, R., & Bercovici, S. The cloak of competence: Years later. *American Journal of Mental Deficiency*, 1976, *80*, 485-497.

Federal Register, Chapter V, Part 525: Employment of handicapped clients in sheltered workshops. May 17, 1974, 17509.

George, H. *Progress and poverty*, 1881, *1* (3), 50.

Gold, M. W. *Remarks on goal-related research needs in the education of the severely handicapped.* Conference on Research Needs Related to the Education of the Severely Handicapped. Princeton, New Jersey: Educational Testing Service and Bureau of Education for the Handicapped, 1975.

Gold, M. W. Task Analysis of a complex assembly task by the retarded blind. *Exceptional Children*, 1976, *43* (20), 78-84.

Greenleigh Associates, Inc. *The role of the sheltered workshop in the rehabilitation of the severely handicapped.* Report to the Department of Health, Education and Welfare, Rehabilitation Services Administration, New York, 1975.

Horner, R. H., & Bellamy, G. T. Stimulus control within operant chains: Variables affecting the vocational performance of one severely retarded worker. Unpublished paper, University of Oregon Center on Human Development, 1978.

Jacobs, J. W. Gleaning: Sheltered employment for retarded adults in rural areas. *Mental Retardation*, 1978, *16*, 168-222.

Jacobs, J. W. Retarded persons as gleaners. *Mental Retardation*, 1976, *14*, 42-43.

Karan, O., Weham, P., Renzaglia, A., & Schutz, R. *Habilitation practices with the severely developmentally disabled.* Madison, Wisconsin: University of Wisconsin, Waisman Center on Mental Retardation and Human Development, 1976.

Kolstoe, O., & Frey, R. *A high school work study program for mentally subnormal students.* Carbondale, Illinois: Southern Illinois University Press, 1965.

Kosmo, R., & Pritchard, K. Disincentives to employment for disabled and dependent populations. Unpublished paper, Rehabilitation Facilities of Wisconsin, Madison, Wisconsin, undated.

Levithan, S., & Taggart, R. *Jobs for the disabled.* Baltimore: Johns Hopkins University Press, 1977.

Lynch, K., & Graber, P. *Survey of Michigan rehabilative facilities: Implications for the developmentally disabled.* Research Monograph, ISMRRD, University of Michigan, Ann Arbor, Michigan, 1977.

Martin, G., & Pallotta, A. Behavior modification in sheltered workshops and community group homes for the retarded: Current status and future considerations. In G. Hamerlynck (Ed.), *Applied behavior analysis techniques for the developmentally disabled.* New York: Brunner Mazel, 1978.

Olshanski, S. Changing vocational behavior through normalization. In W. Wolfensberger (Ed.), *The principle of normalization in human services.* Downsview, Toronto, Canada: National Institute on Mental Retardation, York University Campus, 1972.

Olshanski, S. Job placement: Another discovery. *Rehabilitation Literature*, 1977, *38*(4), 116-117.

Pomerantz, D. J., & Marholin, D. Vocational habilitation: A time for change. In E. Sontag (Ed.), *Educational programming for the severely and profoundly handicapped.* Reston, Virgina: The Council for Exceptional Children, 1977.

Sowers, J., Thompson, L., & Connis, R. The food service vocational training program. In T. Bellamy, G. O'Connor and O. Karan (Eds.), *Vocational rehabilitation of severely handicapped persons: Contemporary service strategies.* Baltimore: University Park Press, 1979.

Turkel, S. *Working.* New York: Pantheon Press, 1972.

U. S. Department of Labor, *Sheltered workshop study.* Washington D. C.: Author, 1977.

Whitehead, C. Sheltered workshops in the decade ahead. In G. Bellamy, G. O'Connor, and O. Karan (Eds.) *Vocational rehabilitation of severely handicapped persons: Contemporary service strategies.* Baltimore, Maryland: University Park Press, 1979.

Wolfensberger, W. (Ed.). *The principle of normalization in human services.* Downsview, Toronto, Canada: National Institute on Mental Retardation, York University Campus, 1972.

A Strategy for Programming Vocational Skills for Severely Handicapped Youth

G. Thomas Bellamy

Darla J. Wilson

Ellen Adler

James Y. Clarke

Preparation of the chapter was funded in part by a grant from the Bureau of Education for the Handicapped, Research Contract #300-78-0346.

This chapter is to be published in Exceptional Education Quarterly, in press. Reprinted with permission.

A Strategy for Programming Vocational Skills for Severely Handicapped Youth

During the past few years there has been a growing professional commitment to vocational preparation of severely handicapped students in secondary programs (Brown, Bellamy, & Sontag, 1971; Sontag, Smith, & Certo, 1977; Fialkowski vs. Pittenger, 1977). However, this consensus about the importance of vocational skill development has not been matched with coordinated efforts to develop longitudinal curriculum sequences. As a result, vocational preparation opportunities are only intermittently available to severely handicapped students and reflect a variety of conflicting theoretical, practical and ideological perspectives.

The purpose of this paper is to propose one strategy for defining and sequencing vocational objectives in secondary programs for severely handicapped students. Our intent is to describe for professional discussion a set of assumptions and procedures that are now being applied to develop instructional sequences.

EDUCATIONAL CONTEXT FOR VOCATIONAL PREPARATION

Vocational preparation of severely handicapped students is consistent with two fundamental characteristics of education in America: a future orientation and a social change perspective. The future orientation of educational programs is apparent in efforts to prepare students for adult roles as citizens, providers, and consumers. Development of skills for predicted vocational environments has long been a part of this future-orientation. For example, Benjamin Franklin lamented in Poor Richard's Almanac that, "Much of the learning in use is of no great use," and suggested that educational emphasis be placed on ... "the several professions for which [students] are intended". Such a utilitarian approach to education was common in early American schools (Cremin, 1970), and despite ongoing debate about the role of education since that time, training for future vocational environments has remained an important focus. Reaffirmation of the importance of such a future orientation was provided in Brown vs. Board (1954) the landmark civil rights decision: "Today (education) is a principal instrument ... in preparing [the child] for later professional training, and in helping him (sic) to adjust normally to his (sic) environment. In these days it is doubtful that any child may reasonably be expected to succeed in life if he (sic) is denied the opportunity to an education." Most secondary students now have access to instruction in several skills related to future work opportunities both in classes and in school sponsored community programs (Evans, 1971).

A second general characteristic of American schools has been an emphasis on social change. Not only is education a means through which students are prepared for future environments, it is also a tool through which future society is shaped. By raising the level of competence of specific groups or by increasing the competence of all students in specific content areas, public schools may have significant effects on future environments and opportunities. (Spring, 1976; Postman & Weingartner, 1969).

IMPLICATIONS FOR EDUCATION OF SEVERELY HANDICAPPED

As severely handicapped students participate in public school programs, these two general characteristics of American education suggest several value considerations for designing instructional programs. First, the opportunity to learn skills that will be relevant to future vocational functioning should be considered central to secondary instruction. These opportunities should be based on a realistic assessment of potential vocational environments and should reflect the wide range of skill areas to which other students are exposed. In addition to providing skills for likely vocational opportunities, schools should attempt to develop vocational skills which might increase the range of eventual vocational options for severely handicapped students. That is, by teaching vocational skills not typically associated with severely handicapped individuals, the schools may be able to expand future opportunities for their students.

These value considerations imply a two-step process for selecting vocationally-related instructional objectives for severely handicapped students. The first step involves prediction of the vocational environments to which students may have access as adults. The predicted environments then can be analyzed to identify both the skills which could improve an individual's performance, and the skills which might improve the predicted vocational options. The second step involves programming, or sequencing vocationally-related objectives, so that the skills identified in the predictions can be taught systematically. This process involves two modifications of traditional task analysis approached to defining instructional objectives for severely handicapped students (e.g., Williams, 1975; Wilcox, 1978). Our emphasis on prediction of future environments necessitates projections about behavioral requirements for competent performance, in addition to direct observation and analysis of performance environments. Second, our emphasis on programming for generalized skills necessitates the application of several rules in sequencing instructional objectives.

PREDICTION OF FUTURE VOCATIONAL ENVIRONMENTS

One way to predict the future is to assume things will remain the same. If this prediction strategy is selected, instructional objectives for vocational preparation could be defined by analyzing current vocational opportunites. Several recent surveys suggest that, at present, severely handicapped adults seldom have access to remunerative vocational environments (Greenleigh Associates, 1975; Urban Institute, 1975; Lynch and Graber, 1977). The consistent finding of these studies was that most adults with moderate or severe retardation and multiple handicaps do not receive services from vocational rehabilitation agencies and are not included in community based sheltered workshops. Rather, they may be enrolled in developmental centers or day activity programs, where a daily regimen of educational and recreational activities presumably prepares them for vocational programs.

Therefore, if currently available work opportunities for severely handicapped adults are assumed to be representative of future opportunities for severely handicapped children and youth, one of three conclusions must be drawn: (a) vocational preparation is irrelevant for these students, since they are unlikely to have access to vocational environments; (b) vocational preparation should be provided only in specific areas of the country where adult service opportunities are considerably above the national average; or (c) vocational preparation should be provided to severely handicapped students in the hope that teaching vocational skills to large numbers of students will be an effective method of changing predicted work opportunities. It is our contention that the premise on which these three options are based is incorrect, and that the discouraging current lack of vocational options is not necessarily predictive of future work opportunities for severely handicapped people.

Predicted Access to Vocational Environments

Despite the pessimistic results of national sheltered workshop surveys, there are several reasons to expect that severely handicapped students now in secondary programs may have access to vocational environments as adults. Recent legislation and litigation, advances in treatment technology, the economics of service, and professional values all suggest that work opportunities may increase.

Legal basis for expecting change. The Rehabilitation Act of 1973 extended to handicapped individuals the civil rights guarantees won by the political activism of blacks and women. This legislation provides a clear basis for challenging any arbitrary exclusion of severely handicapped individuals from vocational services and opportunities. Laski (in press) argues that this legislation and a variety of related court orders provide a convincing basis to argue for an entitlement to vocational and other community services. Whether or not such an entitlement can be established and implemented, much of the legal foundation for extending vocational opportunities to severely handicapped adults is now in place.

Effects of technological advances. The developing sophistication of behavioral and bioengineering techniques is rapidly reducing the logical basis on which job discrimination against the severely handicapped can be justified. No longer can it be assumed that low scores on intelligence tests automatically preclude the acquisition and performance of difficult vocational skills (Bellamy, Horner, Inman, 1979; Gold, 1976). Nor can it be assumed that formerly debilitating physical impairments automatically prevent vocational success. As training, supervision and engineering technologies continue to develop, and as severely handicapped people continue to demonstrate their competence, the right to equal access to vocational preparation and to fair employment consideration should become firmly established.

Economics of service. Several authors have argued that the cost of providing vocational habilitation services to severely handicapped individuals is so high that service priority should be given to less severely handicapped (and, therefore, less costly) clients (Conley, 1973; Kolstoe & Frey, 1965; Levitan & Taggart, 1977). Yet, the cost of not providing vocational services is also high, and it is rising rapidly. The services now provided to severely handicapped individuals through federal and state services and subsidies result in high lifetime public support costs for severely handicapped people. Significant reduction in this anticipated cost could result if partial or full employment helped to defray a portion of the costs associated with personal support, residential and medical care, and daily programming.

Professional values. For some time access to normal opportunities and lifestyles has been considered a central objective in service programs for individuals with handicaps (Nirje, 1969; Wolfensberger, 1972). It is a natural extension of this philosophy that severely handicapped people should have work opportunities. Work is a normal and respected part of most adult living in our society (Turkel, 1972), and as such represents an important objective in service programs (Bellamy, Horner, & Inman, 1979; Gold, 1973).

On the basis of these legal, technological, economic and idealogical considerations it seems reasonable to predict that many severely handicapped individuals will have access to vocational environments as adults. If this is true, then providing vocational instruction in secondary programs may be more than simply a laudable attempt to change future opportunities. Vocational instruction may also be a logical component of future oriented instruction. The content of that instruction, however, must rely on additional predictions. To select specific vocationally related instructional objectives, it is first necessary to identify major characteristics and skill requirements of the work environments to which severely handicapped students may have access.

Characteristics of Future Work Environments

Three projections about future employment opportunities are relevant to vocational preparation of severely handicapped individuals: the structure of the labor force, anticipated manpower needs, and projected job options.

The structure of the labor force reflects both population and social trends. Demographic data suggest that, because of high birth rates in the 1940's and 1950's, the number of working age adults in the United States will increase at least until 1990 (Business Week, 1978). Even greater expansion in the labor force can be predicted during this time as a larger proportion of youth and women seek and obtain work. A similar increase in the size of the labor force during the last few years has not been matched by an increase in the number of jobs. Therefore, a very competitive job market has resulted, and this can be expected to continue. Vocational success in such a market seems more likely for individuals whose skills are relevant to a variety of potential jobs.

Accompanying this competitive labor force is a gradual shift in manpower needs. The number of jobs involved in unskilled labor and in production of goods has been declining for some time and this trend can well be projected during the next decade. Computer technology has eliminated many routine clerical jobs (Evans, 1971) and is likely to affect an increasing number of assembly processes (Watson, 1976). As high-volume, unskilled work is increasingly automated, available industrial jobs probably will require either technical skills or the ability to perform a wide variety of low-volume tasks. Paralleling the automation of production processes has been an increase in service and leisure occupations, which require the employee to perform several tasks in several different settings. Vocational success in either unskilled labor or service occupations would seem to require that employees have flexible work skills, be able to benefit rapidly from instructions, and perform reliably across changing tasks in many environments.

It also seems reasonable to expect that the structure of available job options will change as a result of increasing advocacy for the civil rights of people with handicaps. Few work opportunities currently exist beyond sheltered workshops and individual competitive employment. In sheltered workshops wages are often trivially low and work may be available only intermittently. Whitehead (in press) notes that the average hourly wage for all retarded adults in sheltered workshops is 81 cents and the total monthly wage for these individuals averages only $41.00. Individual competitive employment, on the other hand, is characterized by higher wages, but requires the individual to be independent and competent in a variety of skills indirectly related to the job (e.g., transportation, eating, time telling, self care, social and leisure skills, banking, money, etc.). Further, employment for many handicapped individuals has been only temporary. Frequent job changes, periods of unemployment and intermittent dependence on service agencies have been well documented (Edgerton & Bercovici, 1976). It is our prediction that an increasing array of intermediate employment opportunities will be developed to provide job alternatives that bridge the gap between sheltered and competitive employment. The structured employment model described by Horner and Bellamy (1979) and the workshops within industry approach suggested by DuRand and Neufeldt (1975) are but two vocational options that would allow severely handicapped individuals to be highly productive while providing ongoing personal support, supervision and retraining. It would be possible in these intermediate work environments for handicapped individuals to earn wages applying specific vocational skills before other skill areas less directly related to the job had been mastered.

Skill Implications

These projections suggest that future vocational environments for severely handicapped people will be highly competitive, will require frequent job or task changes and will allow some individual workers to be productive despite deficits in non-vocational skills. If the projections are accurate, there are several implications for vocational instruction of severely handicapped adolescents. First, it is critically important to develop flexible job competencies. That is, instruction should be designed to teach generalized skills that can be applied in a variety of job contexts. Second, instruction would develop skills in a number of different kinds of jobs, to increase the number of positions for which an individual can compete. Finally, if the civil rights of handicapped individuals are reflected in new employment options, the variety of personal and social skills now seen as pre-requisite to work may not completely determine an individual's access to job opportunities. Thus, secondary instruction should emphasize skills specifically required on various jobs as well as personal and social skills that are indirectly related to vocational competence.

PROGRAMMING FOR VOCATIONAL INSTRUCTION

The purpose of programming is to identify and sequence intermediate instructional objectives so that students reach long range educational goals. Programming for vocational instruction with severely handicapped adolescents involves analysis of skill requirements in predicted vocational environments and development of a sequence of tasks through which those skills can be taught.

Several strategies are available for identifying tasks and objectives for vocational instruction. Among these are career awareness instruction, simulated workshop programs and training for specific job descriptions. Career awareness instruction typically is designed to enhance an individual's verbal skill repertoire and "awareness" of various job options. Pictures, films, stories, and visits to work environments allow students to observe skill requirements and potential contingencies of various jobs. This method allows for considerable variety in the job options explored, but the responses required of students during instruction seldom correspond to actual job requirements.

A second strategy for selecting objectives for vocational instruction has involved development of simulated workshops within public school settings (Brown, Bellamy, & Sontag, 1971; Lynch, in press). These workshops typically provide instruction on specific work tasks and give students the opportunity to perform those tasks over time. Students thus are required to make realistic vocational responses and experience contingencies for developing work rate and other appropriate work behaviors. However, the specific tasks used in the classroom or school workshop setting in which instruction takes place may differ considerably from those that will be available in future work situations.

Many of these potential difficulties are avoided by intensive training on specific jobs which are performed in community work situations as part of the instructional process. For example, Belmore and Brown (1978) provide an extensive list of skills needed for successful performance of a dishwashing task in a community job setting. Using such a list as a basis for identifying objectives in vocational instruction insures that the variety of supporting skills required for community based competence will be taught, but it may limit job skill training to a single occupational area.

The general case programming approach suggested in this paper combines many features of each of these strategies. Because of our predictions about future vocational skill requirements, we have chosen a strategy which maximizes the number of job areas to which an individual is exposed and allows the student to perform responses in instruction which are similar to those required on the job. This approach extends the detailed environmental analysis suggested by Belmore and Brown (1978) to multiple tasks and job settings to determine common requirements of several predicted work opportunities.

Our suggestions for general case programming of vocational skills are based on the work of Becker, Engelmann and Thomas (1975), who describe a process for sequencing elementary academic skills. Basic to this strategy is the belief that important instructional objectives are not just single stimulus and response relationships, but rather classes of responses which are controlled by sets of stimuli. In a work setting, for example, a desired product or result is seldom achieved on separate occasions by performing exactly the same responses. Instead, the behaviors involved may differ in topography, sequence, intensity, or timing from one opportunity to the next. The class of behaviors to be taught includes all those variations which achieve the desired result. Similar variability is usually present in the stimuli present in work situations. The required work behaviors, if they are to be applied successfully, must be under control not of a single stimulus array, but any of the combinations of naturally occurring stimuli which indicate that the behavior is appropriate. The purpose of general case programming is to design a series of tasks so that essential characteristics of a set of stimulus classes in natural work environments gain stimulus control over a class of responses, regardless of irrelevant variations within either of the classes. (For a more detailed discussion the reader is referred to Horner & Bellamy, 1978.)

General Case Programming Strategy

General case programming involves a logical analysis of long range educational goals so that instructional objectives can be sequenced in accordance with available research on learning and teaching. The strategy involves the following steps: (a) identify component skill areas; (b) determine the behavioral and stimulus control requirements; (c) sequence instructional content; (d) design instructional tasks.

Identify component skill areas. Vocational preparation programs often rely on clustering work opportunities into broad areas as agriculture, manufacturing, service, and graphic arts. As is illustrated in Figure 1, each area then is sub-divided until a list of specific job descriptions is defined. Such a list of jobs then provides the basis for a variety of vocational preparation programs. Instructional objectives in these programs have high apparent validity because they seem directly related to specific job vacancies. However, the job descriptions on which these objectives are based often do not provide an accurate reflection of the skill requirements of any given job over time. Therefore, basing instructional objectives on specific job descriptions may be inadequate if students are prepared for future work environments in which frequent job changes and rapidly changing tasks with jobs are predicted.

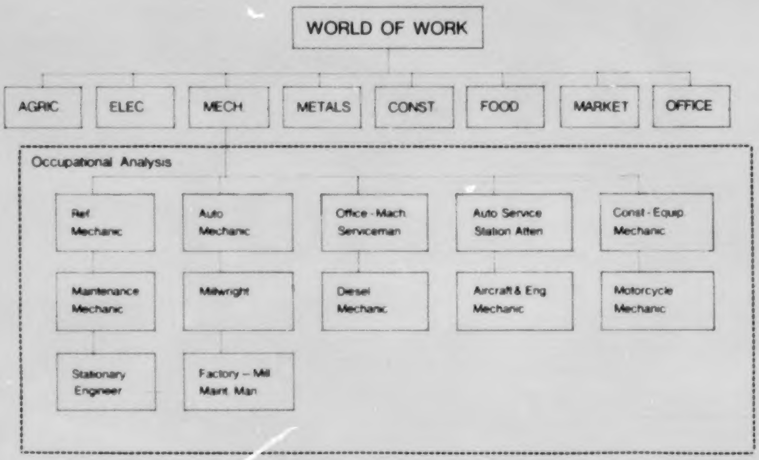


FIGURE 1

Illustration of Division of the World of Work into Job Descriptions

An alternative approach to identifying teachable components from the world of work is to focus on skill areas, or operations (Becker, Engelmann, & Thomas, 1975; Wright & Jensen, 1976). In our analysis a skill area is a response or set of responses which achieves a defined functional effect when it is performed in the appropriate situation. For example, a skill area could be defined as "mechanical fastening with tools," "packaging," "placement of retaining rings" or "repair of flat tires." In each case a particular effect is achieved when a response is performed correctly in the appropriate stimulus situation. Figure 2 illustrates one division of work opportunities into skill areas.

Three instructionally-related advantages result from dividing the world or work into skill areas rather than job descriptions. First, the resulting approach to education is consistent with many recommendations in vocational education literature. Our use of the term skill area closely parallels the terms "process" and "operation" in industrial literature (Wright & Jensen, 1976), and our emphasis on teaching skill areas rather than job descriptions is consistent with the importance vocational educators have attached to teaching processes rather than specific products. Second, dividing the world of work into skill areas allows for the development of flexible vocational skills. Any single skill area could be one component of several jobs. Therefore, combinations of skill areas could qualify an individual for several different positions. A final advantage of this division of the world of work is its compatibility with programming methods. Our definition of skill areas is compatible with the definition of "operation" in the programming strategy presented by Becker, Engelmann and Thomas (1975). This consistency provides the basis for applying a variety of teaching and task sequencing strategies from educated research.

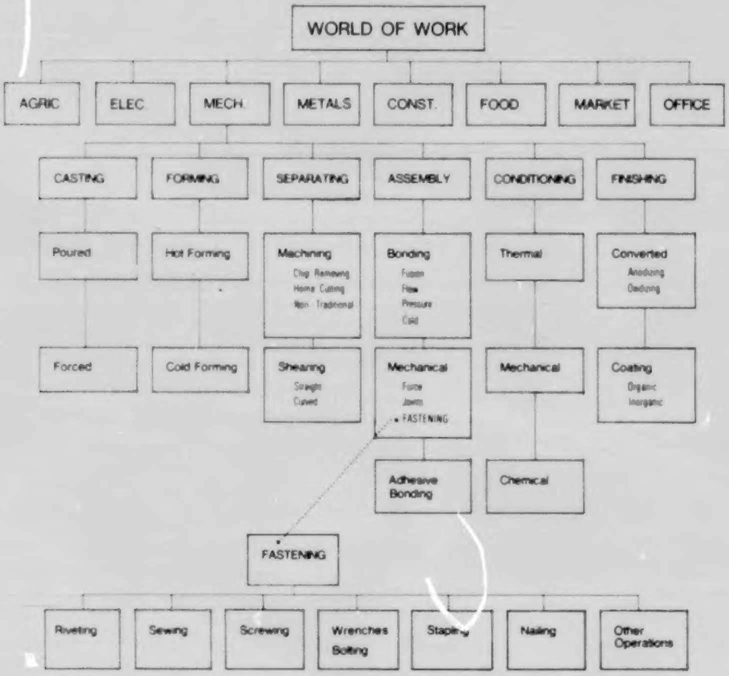


FIGURE 2

Illustration of the World of Work Divided into Skill Areas

Determine behavioral and stimulus control requirements. After a skill area has been identified the next step in general case programming is to identify the behavioral and stimulus control requirements of the skill. This involves (a) identifying the behavior or processes required to achieve the defined effect; and (b) identifying the stimulus features that should and should not control the behaviors if the functional effect is to be achieved.

For example, in the skill area, "using a screwdriver to tighten a screw" the functional effect is the situation when a previously loose screw is tightened. The behaviors required to achieve this effect include seating the tip of the screwdriver in the screw, maintaining perpendicular placement, exerting pressure, rotating in a clockwise position and stopping when the screw is tight. It is important to note, however, that the specific movements involved in each of these critical behaviors may vary widely. The way in which the screwdriver is held, the location of the screw in space, the force required in turning and the length of the screwdriver shaft all affect the exact topography and sequence of movements which are necessary to achieve the defined functional effect. For this reason behaviors involved in any skill area can be usefully conceptualized as involving both essential and non-essential features. Variation or omission of essential features results in failure to obtain the functional effect. Non-essential features may vary without affecting task completion. By analyzing the natural job environment it is possible to identify the essential features of the work behavior and the range of variability in non-essential aspects of the response. Instructional objectives can then be defined so that essential response features are performed despite expected variation in irrelevant response characteristics. (For a more detailed analysis of essential and non-essential features of vocational tasks, the reader is referred to Horner & Bellamy, 1978, and to Bellamy, Horner & Inman, 1979.)

Of course, competence in achieving desired effects in work situations involves not only performance of appropriate behaviors but also performance of those behaviors at appropriate times and places. Thus, the natural situation in which skill areas are performed must be analyzed to identify stimulus control requirements as well as behavioral requirements. Jobs, like all stimulus situations, contain literally hundreds of features to which an individual could attend and respond. Only a few of these typically are essential for determining when a work behavior should or should not be performed. Other stimuli in the environment may frequently change but should not affect whether or not work behavior occurs. Analysis of performance environments to determine stimulus control requirements involves identification of the essential features of the environment which are relevant for job performance and identification of irrelevant features of the environment that vary and that should not affect job performance.

Defining the stimulus control requirements of a skill area in this way, allows deliberate programming for generalization. Instructional objectives and tasks can be designed to insure that the essential responses are performed across the range of nonessential variation in behavioral components, so that critical features of the stimulus situation occasion the response despite irrelevant variation in other aspects of the performance environment. The predicted range of irrelevant variability in stimulus and response components provides guidelines for the range of variability to be included in instructional tasks.

Select sequencing strategies. The third step in applying general case programming to vocational instruction is to define a sequence of instructional tasks through which students can acquire the competencies required by a skill area. These tasks provide the intermediate objectives for instruction in the skill area and should be defined so that teaching is both effective and efficient. Research on learning and teaching has provided several guidelines for establishing sequences which (a) teach required responses, (b) bring these under the control of appropriate stimulus characteristics; and (c) insure performance across irrelevant variation in stimulus and response characteristics.

Sequencing strategies for teaching responses typically involve application of research on shaping and chaining of operant behavior (Skinner, 1953). For example, a variety of task sequences are now available for teaching self-care skills by sequencing individual responses so these can be learned in a forward and reverse order (e.g., Snell, 1978). Other strategies relevant to response development include provision for modeling and physical guidance, massed trials, and sequential introduction of multiple response components (Bellamy, Horner & Inman, 1979).

A wealth of research literature is available to assist in sequencing decisions when the student is expected to learn differential responding to different situations (e.g., Clark, 1971; Zeaman & House, 1963). These include, for example, beginning with simple tasks and requiring progressively more difficult discriminations; developing conjunctive stimulus control by establishing stimulus control first with single stimulus dimensions, then sequentially adding other relevant dimensions; and teaching exceptions to any rule only after the student has learned the general case.

Still other sequencing strategies apply when the objective of instruction involves maintaining performance of a skill despite variation in stimulus and response components. These include systematically sampling the range of expected variation and insuring performance in each setting, and exposing students to extreme cases of expected variation as well as cases within the range, so that students are required to interpolate but not extrapolate from previous experience (e.g., Becker, Engelmann & Thomas, 1975; Stokes & Baer, 1977).

Of course, all vocational skill areas necessitate the application of sequencing strategies for acquisition of responses, development of stimulus control and performance despite variation in stimulus and response requirements. Selecting appropriate sequencing strategies thus involves combining the three types of strategies for achieving each of these results. The relative emphasis on each type of strategy should reflect the programmer's view of the difficulty of the corresponding skill requirement.

Design instruction tasks. The final step in the general case programming strategy is to identify instructional tasks which apply the selected sequencing strategies. An instructional task is a set of materials and/or instructions, together with other situational characteristics, to which the student is expected to respond. A defined criterion for acceptable responding allows for objective monitoring of student progress on instructional tasks.

SUMMARY

Although there is a growing professional commitment to vocational preparation of severely handicapped students, few curricula are currently available to assist in this effort. A strategy for curriculum development was outlined in this chapter that derives from projections about future work opportunities.

Our projection is that severely handicapped students will have access to vocational opportunities, but that a highly competitive labor force and rapidly changing job requirements will necessitate sophisticated and flexible work skills. A strategy for selecting vocationally-related instructional objectives was suggested that appears to fit these projections. The general case programming strategy is designed for teaching generalized skills which can be applied in a variety of job contexts. The strategy involves identifying vocational skill areas, defining response and stimulus control requirements of the skill areas, selecting appropriate strategies and designing instructional tasks. Applications of this and other curriculum development strategies now seems critical if severely handicapped students are to capitalize on vocational opportunities we believe can be available during their adult years.

References

Americans Change. Business Week, 1978, 2522, 64-77.

Becker, W., Engelmann, S., & Thomas, D. Teaching 2: Cognitive learning and instruction. Chicago: Science Research Associates, 1975.

Bellamy, G. T., Horner, R. H., & Inman, D. P. (Eds.). Habilitation of severely and profoundly retarded adults, Vol. II. Eugene, Oregon: University of Oregon, Center on Human Development, 1979.

Belmore, K. J., & Brown, L. A job skill inventory strategy designed for severely handicapped potential workers. In N. Haring and D. Bricker (Eds.), Teaching the severely handicapped, Vol. III. Columbus, Ohio: Special Press, 1978.

Brown, L., Bellamy, T., & Sontag, E. (Eds.). The development and implementation of a public school prevocational training *program for trainable retarded and severely emotionally disturbed children.* Madison Public Schools, Madison, Wisconsin, August, 1971.

Brown vs. Board of Education of Topeka, 347 U.S. 483, 74 S. Ct. 689 (1954).

Clark, D. C. Teaching concepts in the classroom: A set of teaching prescriptions derived from experimental research. *Journal of Educational Psychology*, 1971, *63*(3), 253-278.

Conley, R. *The economics of mental retardation*. Baltimore: Johns Hopkins Press, 1973.

Cremin, A. American Education: *The colonial experience 1607-1783*. New York: Harper and Row, 1970.

DuRand, J., & Neufeldt, A. H. *Comprehensive vocational service systems*. Downsview, Ontario, Canada: National Institute on Mental Retardation, 1975.

Edgerton, R., & Bercovici, S. The cloak of competence: Years later. *American Journal of Mental Deficiency*, 1976, *80*, 485-497.

Evans, N. *Foundations of vocational education*. Columbus, Ohio: Charles E. Merrill, 1971.

Fialkowski vs. Pittenger. Civil Action No. 74-2262 (Eastern District of Pennsylvania, 1977).

Gold, M. Research on the vocational habilitation of the retarded: The present, the future. In N. Ellis (Ed.), *International review of research in mental retardation* (Vol. 6). New York: Academic Press, 1973.

Gold, M. Task analysis of a complex assembly task by the retarded blind. *Exceptional Children*, 1976, *43*(20), 78-84.

Greenleigh Associates, Inc. *The role of the sheltered workshop in the rehabilitation of the severely handicapped*. Report to the Department of Health, Education and Welfare, Rehabilitation Services Administration, New York, 1975.

Horner, R. H., & Bellamy, G. T. Structured employment: Productivity and productive capacity. In G. Bellamy, G. O'Connor and O. Karan (Eds.), *Vocational rehabilitation of severely handicapped persons: Contemporary service strategies*, Baltimore: University Park Press, 1979.

Horner, R. H., & Bellamy, G. T. A conceptual analysis of vocational training with the severely retarded. In M. Snell (Ed.), *Systematic instruction of the moderately, severely, and profoundly handicapped*, Columbus, Ohio: Charles E. Merrill, 1978.

Kolstoe, O., & Frey, R. *A high school work study program for mentally subnormal students*. Carbondale, Illinois: Southern Illinois University Press, 1965.

Laski, F. J. Legal strategies to secure entitlement to services for severely handicapped persons. In G. Bellamy, G. O'Connor, and O. Karan, *Vocational rehabilitation of severely handicapped* persons: Contemporary service strategies, Baltimore: University Park Press, 1979

Levitan, S., & Taggart, R. *Jobs for the disabled*. Baltimore: Johns Hopkins University Press, 1977.

Lynch, K. P. Toward a skill-oriented pre-vocational program for trainable and severely mentally impaired students. In G. Bellamy, G. O'Connor, and O. Karan (Eds.), *Vocational rehabilitation of severely handicapped persons: Contemporary service strategies*. Baltimore: University Park Press, 1979.

Lynch, K., & Graber, P. *Survey of Michigan rehabilitation facilities: Implications for the developmentally disabled.* Research Monograph, ISMRRD, University of Michigan, Ann Arbor, Michigan, 1977.

Nirje, B. The normalization principle and its human management implications. In R. Kugel and W. Wolfensberger (Eds.), *Changing patterns of residential services for the mentally retarded.* Washington, D. C.: President's Committee on Mental Retardation, 1969, 179-195

Postman, N., & Weingartner, C. *Teaching as a subversive activity.* New York: Dell, 1969.

Skinner, B. F. *Science and human behavior.* New York: Free Press, 1953.

Snell, M. E. Self care skills. In M. Snell (Ed.), *Systematic instruction of the moderately and severely handicapped.* Columbus, Ohio: Charles E. Merrill, 1978.

Sontag, E., Smith, J., & Certo, N. *Educational programming for the severely and profoundly handicapped.* Division on Mental Retardation, the Council for Exceptional Children, Reston, Virginia, 1977.

Spring, J. *The sorting machine: National educational policy since 1945.* New York: David McKay Company, Inc., 1976.

Stokes, T. F., & Baer, D. M. An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 1977, *10*, 349-367.

Turkel, S. *Working.* New York: Pantheon Press, 1972.

Urban Institute. *Report of the comprehensive service needs study.* Washington, D. C.: Urban Institute, 1975.

Watson, Paul C. *A multidimensional system analysis of the assembly process as performed by a manipulator, CSDL Report No. P-364.* Charles Stark Draper Laboratory, 1976.

Whitehead, C. W. Sheltered workshops in the decade ahead: Work and wages or welfare. In G. Bellamy, G. O'Connor, and O. Karan (Eds.), *Vocational rehabilitation for developmentally disabled persons: Contemporary service strategies.* Baltimore: University Park Press, 1979.

Wilcox, B. Severe/profound handicapping conditions. In M. Lilly (Ed.), *Children with special needs: A survey of special education.* New York: Holt, 1978.

Williams, W. Procedures of task analysis as related to developing instructional programs for the severely handicapped. In L. Brown, T. Crowner, W. Williams, and R. York (Eds.), *Madison's alternative for zero exclusion: A book of readings*, Vol. V., Madison, Wisconsin: Madison Public Schools, 1975.

Wolfensberger, W. (Ed.). *The principle of normalization in human services.* Downsview, Toronto, Canada: National Institute on Mental Retardation, York University Campus, 1972.

Wright, R., & Jensen, T. *Manufacturing, material processing, management, careers.* South Holland, Illinois: Goodheart-Willcox, 1976.

Zeaman, D., & House, B. The role of attention in retardate discrimination learning. In N. Ellis (Ed.), *Handbook of mental deficiency.* New York: McGraw-Hill, 1963.

Summer Jobs for Vocational Preparation of Moderately and Severely Retarded Adolescents

James Y. Clarke and Lynn M. Greenwood
University of Oregon

David B. Abramowitz
Lane County Employment and Training Program

G. Thomas Bellamy
University of Oregon

This chapter originally appeared in the Journal of the Association for the Severely Handicapped, 1980, 5, 24-37. Reprinted with permission.

Summer Jobs for Vocational Preparation of Moderately and Severely Retarded Adolescents

In vocational preparation, as in other curriculum domains, educators of moderately, severely and profoundly handicapped students have generally avoided techniques that relied on simple exposure to task requirements or classroom verbal exchanges to teach useful skills. Instead, instructional strategies have been developed that require students to learn and perform the specific skills needed in natural living environments, (Brown, Nietupski & Hamre-Nietupski, 1976).

This effort to introduce realistic performance requirements into classroom programs has resulted in at least three strategies for vocational preparation of severely handicapped youth. First, simulated workshops have been developed in several secondary schools to teach the skills required for sheltered employment (e.g., Brown, Bellamy & Sontag, 1971; Lynch, 1979. These settings typically allow for both instruction on specific vocational skills and improvement of production rate and social behaviors required for sustained work. However, the types of skills taught often reflect only the tasks and performance environments of sheltered work settings. The possibility that many severely handicapped individuals could succeed in competitive employment (Cook, Dahl & Gale, 1977) suggests that public school workshops may prepare students for an unduly restricted range of vocational skills.

A second strategy, exemplified by the work of Potter, Biacchi and Richardson (1977) avoids some of the restrictions of a school workshop by simulating several vocational and daily living tasks in portable training environments. The diversity of skills on which a student could receive instruction appears to be increased by this simulation strategy, but little information is now available on how it could be used to develop the sustained performance and concomitant social behaviors required in most jobs.

The third strategy, described by Belmore and Brown (1973) involves using community-based work environments to teach the skills required for a defined job. This approach would appear to increase correspondence between instructional objectives and actual job skill requirements, because teaching occurs in the natural work setting, where the necessity for appropriate social behavior, independent commuting and personal management skills is normally present.

A further improvement in the realism of vocational preparation is now possible because of 1979 changes in the Comprehensive Employment and Training Act (CETA) Program regulations (Federal Register, 1979). Economic eligibility criteria have been removed for people with handicaps, making it possible to provide fully paid work training and experience as a part of job preparation in most communities. Thus, not only can severely handicapped students learn and perform work skills in natural work settings, but they also can work entire days, function as a part of social groups at work, and enjoy the benefits, problems and learning experiences created by access to normal wages.

This paper provides a description and evaluation of one program approach that used CETA work experience funds to train and place moderately and severely handicapped secondary students on summer jobs. The evaluation of the project was made with multiple measures including consumer ratings of satisfaction with the program, participant earnings, documentation of job skills learned and performed and a measure of change in parent's attitudes towards their handicapped child.

PROGRAM METHODS

The Special Needs Program was a joint project between the Lane County, Oregon, Comprehensive Employment and Training Act (CETA) Program and the Specialized Training Program at the University of Oregon. The project was designed around a co-worker concept in which one non-handicapped and one handicapped adolescent formed a working team for a full or part-time community job. CETA staff identified appropriate jobs, selected CETA- eligible non-handicapped co-workers, and paid the salaries of all participants. The Specialized Training Program provided co-worker training and assisted with the job skills training of the handicapped students. The project involved: selection of handicapped participants, selection and training of co-workers, selection and analysis of jobs, and program supervision.

Selection of Workers

The student workers were drawn from five secondary level classrooms for moderately and severely handicapped students in the Eugene/Springfield area. Two students from each classroom were served by the program. Each classroom teacher selected two students who they believed would benefit most from the summer employment experience. With two exceptions, these teacher selections became the participants in the program. Of the ten student participants, eight were labeled moderately retarded and two severely retarded on the basis of diagnostic information in school records.

Selection and Training of Co-workers

Co-workers were selected by program staff from 16 to 21 year old CETA-eligible young people who had career interests in human services. In general the young people selected had little experience in working with handicapped individuals.

The co-workers were trained in groups of two or three for three training sessions. The first session's content was information about handicapping conditions and current trends in the habilitation of severely handicapped individuals. The final two sessions consisted of training modules designed to teach training skills based on applied behavioral analysis research with particular emphasis on the following four skills:

1. **Positive reinforcement**: The delivery of positive consequences contingent on appropriate work behavior for the purpose of training new work skills and maintaining performance of already learned skills.
2. **Use of assistance in training**: The provision of assistance to workers who need to develop a job required response or improve the quality of an existing response. Co-workers learned to use three forms of training assistance: a) verbal instruction, b) modeling a skillful response, and c) physically priming effective work behavior.
3. **Task analysis**: Identification of component tasks of typical community jobs. Co-workers were trained to analyze a job task into a sequence of response components which, if performed, comprised competent performance of that task (O'Neill & Bellamy, 1978).
4. **Data collection**: Verification of the acquisition and performance of work skills by the collection of data on correct and incorrect responding during acquisition of a work skill and data on performance of work skills after a task was learned.

Job Selection and Analysis

Jobs for the student participants were selected from a pool of jobs identified by a CETA job developer for summer youth employment. The jobs were typical of local work for unskilled workers first entering the labor market. Each potential work site was visited by two members of the program staff to define the tasks that comprised the job. In most instances the staff had the work supervisor describe the job and demonstrate essential tasks necessary to perform the work (Davies, 1973).

After the job was analyzed for task components, a co-worker was placed at the site. The co-worker performed the work for a minimum of one week. During this time the co-worker, with the assistance of a staff member, became familiar with the manner in which job tasks were performed. Daily work schedules were developed and difficult job tasks were analyzed into a sequence of response steps to assist in training.

Program Supervision

Supervision procedures were designed to accomplish three objectives: 1) to facilitate the acquisition of appropriate work skills by the handicapped workers; 2) to maximize independent performances of work skills learned; and 3) to monitor appropriate social and general work behavior of both the co-worker and handicapped worker in the work setting. The first of these, facilitating the acquisition of work skills, was accomplished through direct instruction on required skills, using assistance, reinforcement and assessment procedures described by Bellamy, Horner, and Inman (1979). Task analysis data sheets maintained by co-workers provided a visual display of learning progress and identified response steps consistently performed correctly and incorrectly (Bellamy, et al., 1979). This information allowed staff supervisors to assist co-workers in decisions about training procedures that would aid the workers' job skills acquisition.

The second objective, maximizing independent performance of skills, involves efforts to assist the handicapped student meet all job expectations relevant to rate, quantity and quality of work. Job task lists maintained by co-workers allowed staff supervisors information about which job tasks were performed daily, how long each was performed and whether the student worker performed independently or not. These data were useful in making decisions about task retraining and evaluating whether the contract with the work site was an accurate reflection of the type and quantity of work required. Because time standards for work tasks were not available for most jobs, many of the judgments about adequacy of rate and quality of work performed were made subjectively by staff.

The third objective, appropriate social and general work behavior by co-workers and handicapped workers, was monitored in two ways. First, each on-the-job supervisor was asked to complete a short evaluation form each week that asked about her or his satisfaction with the overall program and the social behavior and work performance of both the co-worker and handicapped worker. If this information communicated a perception, by the on-the-job supervisor, that the performance of the student or co-worker was inadequate or less acceptable than the previous week, the staff increased on-site supervision. When a specific social or work behavior deficiency was identified, a structured intervention program was developed to change social or work performance to acceptable levels. The second method of monitoring work and social behavior was through the subjective views of staff during bi-weekly on-site visits. If a problem was identified a variety of counseling and behavior change strategies were employed for remediation of the problem.

CASE STUDIES

Two case studies are included to provide more detailed information on individual work settings and to describe training and work performance on these jobs. The two work situations were typical of jobs in the program.

Worker One

Megan was an 18-year old Down's syndrome woman who was selected for the program from a local secondary class for severely handicapped students. Megan was labeled moderately retarded on the basis of a standardized intelligence test and adaptive behavior test administered by a certified educational psychologist. She has expressive and receptive language skills, can write her name, tell time, and perform all self-help skills independently. Prior to Megan's involvement in this program she had received no vocational training in school or in the community. During the project she lived at home with her parents who were supportive of her involvement in the program but doubtful that she would be able to do well on job related tasks.

Megan's job site was an agency which distributed books, catalogues, and other materials by mail to book stores and libraries throughout the United States. The agency employed four full-time and several part-time people and was located on the third floor of a modern office building. Megan and her co-worker were originally hired to prepare books for mailing. Each book was individually wrapped in both heavy paper and cardboard, and then weighed and labeled for mailing. This task was broken into 28 steps for training purposes (see Figure 1). To perform the task Megan used scissors, a paper cutter, special wrapping tape and a postage scale. Books to be wrapped were of varying sizes, so the job involved several discriminations on length of tape, paper and cardboard.

S^D	RESPONSE
1. Invoice in to be mailed box	Pick up invoice
2. Invoice picked up	Go to shelf area
3. At shelf area	Locate book by title
4. Proper book in hand	Bring book to table
5. Book at table	Pull out sufficient brown paper
6. Sufficient brown paper	Set book on paper
7. Book on paper	Take label off invoice
8. Label off invoice	Put invoice on book
9. Invoice on book	Wrap brown paper around book
10. Brown paper around book	Pull out and cut four 3" pices of tape
11. Four 3" pieces of tape	Tape paper tightly around book
12. Paper taped tightly around book	Put piece of wrapping cardboard under book
13. Cardboard under book	Bend cardboard around book
14. Cardboard bent around book	Cut off excess cardboard
15. Cardboard trimmed	Pull out nine 2" pieces of strapping tape
16. Nine 2" pieces of strapping tape	Shape cardboard for ends
17. Cardboard shaped for ends	Cut cardboard for open ends
18. Open ends cut in cardboard	Tape open ends tightly with strapping tape
19. Open ends taped tightly	Measure out sufficient paper tape
20. Sufficient paper tape	Wet tape or sponge
21. Tape wet	Tape ends of book with paper tape
22. Ends of book taped	wet mailing label
23. Mailing label wet	Place label on book
24. Label in book	Weigh book and write weight on corner of book
25. weight on book	Look at postage rates and write rate on corner of book
26. Amount of postage on corner of book	Stamp postage with appropriate rubber stamp
27. Book appropriately stamped	Put finished book in outgoing box
28. Finished book in outgoing box	Pick up next invoice

FIGURE 1

Task Analysis of Wrapping Book for Mailing

Training began in May for two hours a day after school and initially it took Megan and her co-worker the entire session to complete one book for mailing. The industry standard rate for the task was one book per 20 minutes. During the first training sessions Megan required assistance on 15 of the 28 steps. Two difficult measuring steps were made easier by developing measuring cues to use when pulling out lengths of paper and cutting appropriate sized pieces of cardboard. A staff person worked with the team on training every day for the first two weeks and then checked on the work in progress three times a week thereafter. Megan and her co-worker began working a six-hour day as soon as school was out in June (about four weeks after being placed at the job site).

Megan and her co-worker worked at a large table in the back of the main office work room. They received invoices from a supervisor daily and worked independently in the setting.

By the end of the summer Megan could complete the wrapping task without errors and have the book ready for mailing in 30 minutes. Occasionally Megan required a verbal reminder on two or three of the difficult steps. Megan's supervisor, who monitored her work at these times, reported that she worked competently and at an acceptable rate. Megan rode the bus from her home to a transfer station and from the transfer station to her home independently by the summer's end.

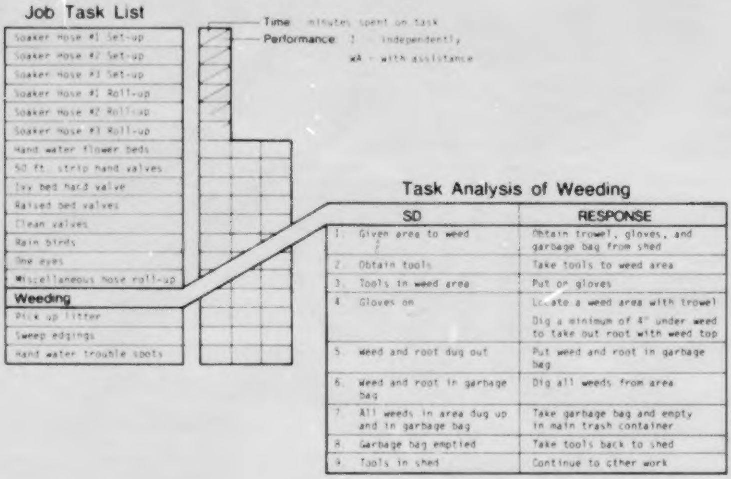
Megan worked six hours a day for 11 weeks during the summer. Her rate of pay was $2.65 per hour and she earned approximately $1,100 for her work.

Worker Two

Al is an 18-year old young man who is non-verbal, physically handicapped, and labeled moderately retarded on the basis of a standardized intelligence test and adaptive behavior test administered by a certified educational psychologist. During the school year he attends a classroom for severely handicapped adolescents in a segregated private school. Al never had a job before.

His job was with the maintenance department at the University of Oregon. The primary task was to water a large area between two buildings (approximately two acres) and some small sections around this area that have no automatic sprinkling system. This task consisted of setting up a variety of sprinklers, moving them periodically to new locations, and rolling up the hoses and putting the equipment away at the end of the work day. Additional job tasks included weeding, cleaning up after edging, policing the area and emptying trash containers.

The method of training for this job was modeling competent performance of job tasks for Al and giving him an opportunity to perform the task. If he had difficulty performing a job task it was analyzed into a sequence of smaller steps for training. Figure 2 illustrates the task components of this job and expands one task, weeding, into a task analysis that was required because of Al's difficulty in the acquisition of this skill.



Job Task List

Soaker Hose #1 Set-up
Soaker Hose #2 Set-up
Soaker Hose #3 Set-up
Soaker Hose #1 Roll-up
Soaker Hose #2 Roll-up
Soaker Hose #3 Roll-up
Hand water flower beds
50 ft. strip hand valves
Ivy bed hand valve
Raised bed valves
Clean valves
Rain birds
One eyes
Miscellaneous hose roll-up
Weeding
Pick up litter
Sweep edgings
Hand water trouble spots

Time - minutes spent on task
Performance I - independently
WA - with assistance

Task Analysis of Weeding

SD	RESPONSE
1. Given area to weed	Obtain trowel, gloves, and garbage bag from shed
2. Obtain tools	Take tools to weed area
3. Tools in weed area	Put on gloves
4. Gloves on	Locate a weed area with trowel Dig a minimum of 4" under weed to take out root with weed top
5. weed and root dug out	Put weed and root in garbage bag
6. Weed and root in garbage bag	Dig all weeds from area
7. All weeds in area dug up and in garbage bag	Take garbage bag and empty in main trash container
8. Garbage bag emptied	Take tools back to shed
9. Tools in shed	Continue to other work

FIGURE 2
Ground Maintenance
A Listing of Job Tasks and a Task Analysis of Weeding

The task list was utilized to record which job elements were performed daily, how much time was spent doing each task, and whether tasks performed were done independently or required co-worker assistance for accurate performance. Figure 3 represents the number of job tasks that Al performed independently across days of work. After 10 days of full-time work Al performed 17 out of 18 job tasks independently.

Al was enthusiastic about his work and performance of tasks was not a problem. As the summer progressed he worked one-hour periods without supervision. The discriminations with which he required regular assistance over the entire summer was when sections of ground were watered sufficiently and others were due for watering.

Al got along well with his fellow workers and supervisors and shared daily lunch and break periods with them. When his co-worker was required to miss a two-day period, Al's fellow workers and supervisor volunteered to take over the co-worker's function and Al was able to continue working full-time. Al's parents reported that he gained tremendously in self-confidence during the summer and was very proud of the paychecks he earned. He left for a family vacation two weeks before the program's termination. He worked 7 1/2 hours a day for eight weeks and earned over $800.00

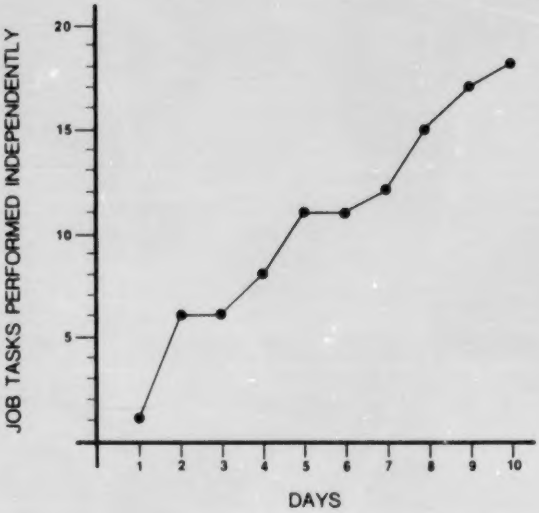


FIGURE 3

Number of Job Tasks Performed Independently by Al after 10 Days of Work

RESULTS AND DISCUSSION

The program was successful in providing both work opportunity and vocational training that allowed nine of the ten handicapped students to perform remunerative and useful work during the summer. Task analyses developed as training tools for teaching job tasks and the daily job task lists provided data which confirm that all participating students acquired and performed job skills. Three measures, participant earnings, parental rating of student change and parental consumer satisfaction ratings, describe the effects of the summer employment program.

Participant Earnings

Each handicapped student earned $2.65 an hour. The nine successful placements worked between eight and eleven weeks during the summer. Three of these workers worked two hours after school in May at their jobs. The total earnings for the nine students were $8,700.

Parental Rating of Student Change

An adaptation of a bipolar adjective checklist originally designed by Becker (1960) for parents and teachers of young children was utilized to measure change in parents' attitudes towards their working child. The device is a semantic differential checklist which consists of pairs of opposite adjectives, with the adjectives in each pair defining the extremes of a seven-point rating scale. The adjective pair represent a range of negative to positive ratings on a variety of attributes of the student workers. Before the program began each parent completed this test (pretest) and after the work experience was complete each parent completed the same test (posttest).

The adaptation of this test has been factor analyzed on a sample of retarded workers by Bellamy and Irvin (1977) and five subscales were empirically determined. Table 1 reports pre and posttest means and standard deviations for the sample of eight student workers on these five subscales. (One parent did not complete the forms.) Note that the more negative a score the less inappropriate or deviant the student is considered; i.e., movement in a negative direction connotes increased social desirability. The magnitude of the treatment effect is expressed in standard deviation units, a technique recommended by House, Glass, McLean, and Walker (1978) when dealing with small sample sizes in quasi-experimental settings. Generally a difference of more than one-third standard deviation is considered educationally significant (Horst, Tallmadge & Wood, 1975). All five differences are educationally significant by this criteria. Gains in the following factors are also statistically significant by one tail-tests. Withdrawn, $t_{(7)} = -1.93$, $P < .05$; Tense, $t_{(7)} = -4.07$, $P < .01$; and Non Comply, $t_{(7)} = -3.05$, $P < .01$. There is thus evidence of both educationally

TABLE 1

Pretest and Posttest Means and Standard Deviations of Adapted Subscales of Becker Sematic Differential for Eight Subjects in a Summer Work Program

	Pretest	Posttest	Magnitude of Effect in SD Units
Withdrawn	$\bar{x}$ = -17.5	$\bar{x}$ = -22	.6 SD
	Sd = 7.56	Sd = 7.27	
Disruptive	$\bar{x}$ = - 6.4	$\bar{x}$ = - 8.25	.4 SD
	Sd = 6.09	Sd = 7.09	
Incompetent	$\bar{x}$ = - 4.88	$\bar{x}$ = - 6.63	.4 SD
	Sd = 5.06	Sd = 3.85	
Tense	$\bar{x}$ = .38	$\bar{x}$ = - 4.13	.7 SD
	Sd = 7.48	Sd = 6.08	
Non-compliant	$\bar{x}$ = - 4.88	$\bar{x}$ = -10.38	1.0 SD
	Sd = 5.11	Sd = 5.88	

and statistically significant improvements in the workers in the areas of social withdrawal, tenseness and non-compliance, as viewed by their parents. Thus parents labeled their childs' behavior as more normative on three of five dimensions measured by the test. These improvements should be interpreted with caution, however, since no untreated control group was used to rule out the possibility that such changes could have resulted from simple maturation, repeated measurements, or other factors.

Consumer Satisfaction: Parental Ratings

It is the view of the program staff that the students valued their work experiences but measurement devices are not presently available that reliably allow mentally handicapped adolescents to judge the quality and effectiveness of programs designed to meet their vocational needs. Therefore, it was felt that another important group in a position to make judgments about the effectiveness of this program were the parents of the handicapped workers. A consumer satisfaction form was completed by eight of the parents. The questions ranged from their rating of the program's performance in teaching job skills to an overall rating of the program's effectiveness. A score of 4 was equivalent to an excellent rating and a score of 1 was a poor rating with 3 and 2 representing an average + and average - rating. The lowest rating of the program was an 18 score for 6 questions or a 3 average. Four of the eight parents scored all six questions as 4's, or excellent ratings of program quality. The average score for each question was 3.7. Overall, the parents considered the program as effectively and professionally providing a valuable vocational experience to their child.

Suggestions for Similar Programs

One handicapped student and co-worker were dropped early in the program. The student had a difficult combination of health and behavior problems. The co-worker had significant personal problems that affected her ability to do her work. Additionally, the program staff were unable to obtain a consistent work activities schedule from the work supervisor or the general cooperation of fellow workers with the program. This unsuccessful placement highlights the two major problem areas faced by the program: negotiating a sufficient and consistent work schedule for each work site and obtaining reliable work performance from the co-workers.

If a program such as this were repeated it is recommended that these two problem areas should be approached with additional program procedures. First, the job selection process should be standardized so that a contract would be made between the program and the job supervisors that carefully defined the nature of duties to be performed, quality standards for the work, the quantity of work guaranteed to be available to the working team daily and what type of supervision from the work site staff would be given to the working pair. Second, because of a high rate of absenteeism, personal problems affecting work performance, and interactional problems with regular work personnel by the co-workers, additional co-worker training is recommended.

Clear standards for absences from work in terms of number of times allowed and mandatory procedures to go through if an absence was necessary should be established. Co-workers should be informed of the importance of separating non-work related personal problems from affecting their work performance. Finally, the distinction between co-worker role and supervisor role should be more clearly established and communicated to co-workers.

CONCLUSION

These results suggest that summer employment for severely and moderately handicapped adolescents is a viable job preparation strategy. The program described was successful in teaching student workers job skills and providing work environments for daily performance of those skills. Parents' attitudes about their child in general and their vocational potential specifically improved significantly over the summer. Higher and improved expectations about their children's working future may result. Parents acting as representatives of the student consumers of this program rated the project as effective in its purpose of finding jobs, training students for a job and having students remuneratively working during the summer.

The design of this program was consistent with the research literature (Stokes, Baer, & Jackson, 1974; Walker & Buckley, 1972) which suggests that teaching handicapped individuals skills in the environments in which those skills are to be performed is an optimal training tactic. The additional advantage of this project was the provision of realistic experiences with the natually occurring benefits and problems of remunerative full-time work.

Clearly, additional research is necessary for optimal delivery of similar programs. Two of the questions of importance are the identification of the exact skills co-workers need to have for effectiveness and the development of data systems that are manageable yet meet the data needs of trainers and program managers.

References

Becker, W. The relationship of factors in parental ratings of self and each other to the behavior of kindergarten children as rated by mothers, fathers, and teachers. Journal of Consulting Psychology, 1960, 24, 507-527.

Bellamy, T., Irvin, L. Utility of an instrument for assessing labels used to describe retarded adults. In T. Bellamy, R. Horner, and D. Inman (Eds.), Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program, Vol. II. Eugene, Oregon: University of Oregon Center on Human Development, 1977.

Bellamy, G. T., Horner, R. H., & Inman, D. P. Vocational habilitation of severely retarded adults. Baltimore: University Park Press, 1979.

Belmore, K., & Brown, L. A job skill inventory strategy for use in a public school vocational training program for severely handicapped potential workers. In L. Brown, N. Certo, K. Belmore, and T. Crowner (Eds.), Madison's alternative for zero exclusion, Vol.VI. Madison Public Schools, Madison, Wisconsin, 1978.

Brown, L., Bellamy, T., & Sontag, E. The development and implementation of a public school prevocational training program for trainable retarded and emotionally disturbed children. Madison Public Schools, Madison, Wisconsin, 1971.

Brown, L., Nietupski, J., & Hamre-Nietupski, S. The criterion of ultimate functioning and public school services for severely handicapped students. Madison, Wisconsin: University of Wisconsin and Madison Public Schools, 1976.

Cook, P., Dahl, P., & Gale, M. Vocational training and placement of the severely handicapped: Vocational opportunities. Palo Alto: The American Institute for Research in the Behavioral Sciences, 1977.

Davies, I. Competency based learning: Technology, management and design. New York: McGraw Hill, 1973.

Federal Register, Part IX: Department of Labor; Comprehensive Employment and Training Act Regulations; Final rule. April 3, 1979.

Horst, D. P., Tallmadge, G. K., & Wood, C. T. A practical guide to measuring project impact on student achievement. Monograph No. 1 on evaluation in education. Washington, D. C.: U. S. Government Printing Office, 1975.

House, E. R., Glass, G. V., McLean, L. D., & Walker, D. W. No simple answer: Critique of the "follow through" evaluation. Harvard Educational Review, 1978.

Lynch, K. Toward a skill-oriented prevocational program for trainable and severely mentally impaired students. In G. Bellamy, G. O'Connor, and O. Karan (Eds.), Vocational habilitation of the developmentally disabled: Contemporary service strategies. Baltimore: University Park Press, 1979.

O'Neill, C., & Bellamy, T. Evaluation of a procedure for teaching saw chain assembly to a severely retarded woman. Mental Retardation, 1978, 16(1), 37-41.

Potter, J. T., Biacchi, A. J., & Richardson, E. A. Simulating real-life situations in a classroom setting: The Montgomery County training modeule. In E. Sontag (Ed.), Educational programming for the severely and profoundly handicapped. Reston, Virginia: Council for Exceptional Children, 1977.

Stokes, T. F., Baer, D. M., & Jackson, R. L. Programming the generalization of a greeting response in four retarded children. Journal of Applied Behavior Analysis, 1974, 7, 599-610.

Walker, H. M., & Buckley, N. K. Programming generalization and maintenance of treatment effects across time and across settings. Journal of Applied Behavior Analysis, 1972, 5, 209-224.

Information Processing for Habilitation

Shawn Boles and J. Daniel Boomer

This chapter was prepared with support from the Bureau of Education for the Handicapped Research Grant #54-P-71693/0.

Information Processing for Habilitation

During the last twenty years we have seen the establishment of two equally strong and positive forces acting to increase the quality of life for citizens who heretofore were relegated to a marginal existence in the bowels of state institutions. The first of these began in the judicial arena and has resulted in a court mandated shift in emphasis from the institution to the community as a source of treatment and support for citizens handicapped by developmental and other disabilities. The second began in the area of educational technology and has developed into a powerful set of programming strategies which provide handicapped people with an ever increasing ability to learn and perform in a fashion that is acceptable to the communities of which they are members.

Concurrent with these two forces, however, has come an increasing awareness of the logistical difficulties encountered in providing mandated habilitation and treatment services in community settings. These difficulties revolve in part about the problems associated with ensuring that service providers have access to the demographic, programmatic, and financial information required to establish and meet their service objectives (Schalock, 1978). This paper reports one attempt to resolve some of those problems by developing a computer-based information system for service providers working with a particular population (severely retarded adults) in a number of community based workshops utilizing the Specialized Training Program model of habilitation. Specifically, this report focuses on the use of a system to provide programmatic information to service personnel for tracking workshop and individual client production performance. Following this description, consideration is given to future development of the system to integrate demographic and financial information with the programmatic data base already established.

Specialized Training Program Information System

User characteristics

The Specialized Training Program (STP) model of habilitation is described in detail elsewhere (e.g., Bellamy, Horner, and Inman,1979 ; Bellamy, Inman, and Horner, 1979), but can be summarized as having an applied operant orientation to worker training coupled with sound business practices in small community based service settings. This model is employed by five subcontract workshops located throughout the western United States, each of which serves approximately fifteen severely retarded workers with a three person staff. These workshops, in addition to employing a common set of training and production management strategies, are also joined together in a management services organization (The Association of Specialized Training Programs) upon which they rely for contract procurement, technical assistance, and information processing services.

In any service delivery system there are multiple users of information regarding the extent to which service objectives are being met. STP workshops are no exception, and it is possible to identify at least six separate categories of users associated with the STP network of workshops who must have programmatic data in order to perform their work activities:

Production Supervisors responsible for scheduling and maintaining production within the workshops require information about the productivity of individual workers on particular tasks. This information is needed to estimate the amount of work that can and is being absorbed by the workshop. They also require such information in order to track the results of individualized programs designed to optimize worker performance in the production setting.

Workshop Directors responsible for overall operation of the workshops must have information about the amount of time workers spend in training on particular tasks in order to forecast and track the preproduction activities associated with contract acceptance. They also require information about worker and workshop earnings stemming from production tasks in order to compile payrolls and forecast production income.

Boards of Directors charged with evaluating workshop activities require information about the aggregate performance of the workshops in terms of overall productivity, proportion of time in production, and gross earnings. This facilitates direction and support of the workshop staff in their endeavors.

Agency personnel such as those working with state Vocational Rehabilitation agencies, the Department of Labor, and accreditation agencies, require regular reports on client progress. Data on work rate, earnings, and time in program provide information for program reimbursement and evaluation purposes.

Research and Policy staff from agencies supporting the development of the STP model (e.g., Department of Health, Education, and Welfare: Developmental Disabilities Office, etc.) require longitudinal information about the effectiveness of the model in terms of proportion of time allocated to training, total workshop productivity, and worker contributions to workshop income.

Sales staff charged with contract procurement and maintenance for the workshops require information about workshop performance on individual contracts in terms of productivity, units produced and error rate. This information is used both for feedback to current customers and as a sales tool for potential contractors.

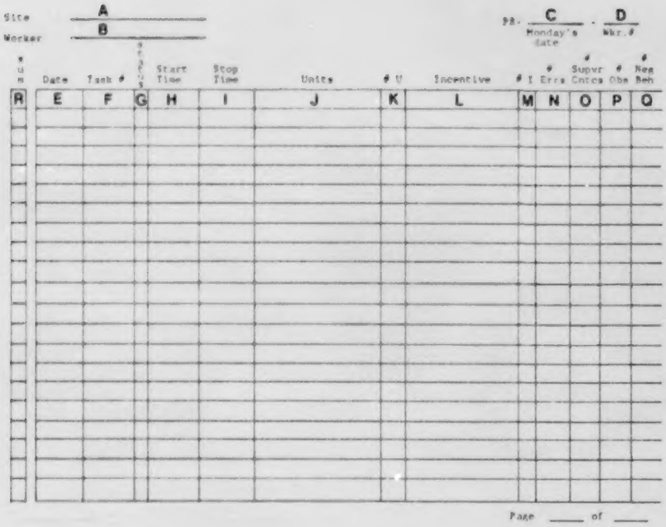
Given the diverse group of information users associated with the workshops and the variety of reporting formats required to meet their needs, a decision was made to develop a computer based system that would maximize the number of ways the basic input data could be analyzed and displayed. At the same time, an equally strong decision was made to ensure that workshop staff who collected the basic input data would not be burdened by the collection process to the point that it would interfere with their other responsibilities. The remainder of this section describes the system resulting from these decisions.

Input

Input for the information system consists of two parts. The first of these is a task form which lists the number of each task currently available to the workforce in a particular workshop, the name of the task, the industrial time standard for production of a single unit of the task, the worker pay per unit produced, and the workshop income per unit produced. Workshop directors fill out this form and submit it to the STP offices at the Center on Human Development at the University of Oregon where it is entered into the University PDP-10 computer. A workshop director can change the task form data whenever s/he desires by simply submitting an updated version of the form containing only the changed data.

The second part of the system input is collected as a regular part of the daily activities of supervision staff who are working in the production and training areas of a workshop. Collection occurs through the use of a production behavior form associated with each worker. This form is shown in Figure 1 and contains space for the following:

a. "site". The name of the STP workshop where the worker is employed.
b. "worker". The name of the worker.
c. "Monday's date". Monday's date for the current work week.
d. "worker #". A two digit number unique to the worker whose name appears in b.
e. "date". Daily date work station activity occurs.
f. "task #". A unique five digit number associated with the task the worker is working on.
g. "status". A one letter code indicating whether the task being worked on is to be considered production, training for production, simulation, or training for simulation.



Site A

Worker B

PB- C - D

Monday's date

Wkr.#

Sum	Date	Task #	Status	Start Time	Stop Time	Units	# U	Incentive	# I	# Errs	# Supvr Cntcs	# Obs	# Neg Beh
R	E	F	G	H	I	J	K	L	M	N	O	P	Q

Page ____ of ____

FIGURE 1

Production Behavior Form

h. "start time". The time of day the worker starts work on the task.

i. "stop time". The time of day the worker stops work on the task.

j. "units". This space is provided for tallying units completed and/or comments on worker behavior.

k. "# units". The number of units produced correctly between the start and stop times.

l. "incentive". This space allows tallying the number of incentives the worker receives.

m. "# I". The number of incentives received between the start and stop times.

n. "# errors". The number of units produced containing one or more errors.

o. "# supervisor cntcs". The number of times the supervisor interacted with the worker.

p. "# obs". This space allows the recording of observations made on any special behavior of interest to the supervisor. (e.g., off-task behavior, screaming, etc.)

q. "# neg beh". The number of times a special behavior of interest occurred.

r. "sum". A check mark entered here indicates that data on this line summarizes previous lines for the same task and status.

Summary lines are entered at varying times for each individual worker with a minimum occurrence of one summary line per worker per task per status per day for each task worked. Since there is a production behavior form at each worker's work station it is easily filled out in conjunction with the normal contact between staff and workers occurring throughout the work day (e.g., quality control checks, parts replenishment, individual program implementation, etc.). Rarely does filling out a line on the form take a staff member more than ten seconds.

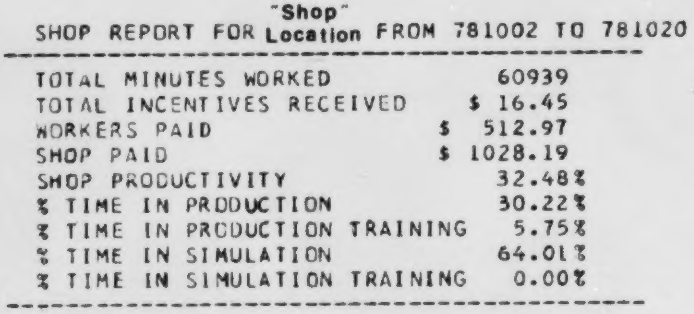
Production behavior forms for a workshop are collected each week and mailed to the STP offices where they are entered into the computer. This usually takes three days for mailing and one day for entry.

Output

Output from the system to the users is in the form of a series of reports produced upon a schedule specified by the users. In order to meet the varying times at which each user might wish to obtain a report an interactive report generation program is set up to produce reports covering from one day to six weeks of input data. In addition, input data older than six weeks are stored on magnetic tape so as to form a longitudinal data base for research and policy purposes.

Reports may be broken down into two basic types: (a) workshop reports that summarize the activity of the entire workforce, and (b) worker reports that deal only with a single worker. Examples of each of these types will be discussed in turn.

Workshop Reports. These reports provide summary information to workshop directors, their boards, and the sales staff. Figure 2 is an example of the first of these reports: the workshop summary. The workshop summary for each workshop shows the time period covered by the report; the workshop location; the total



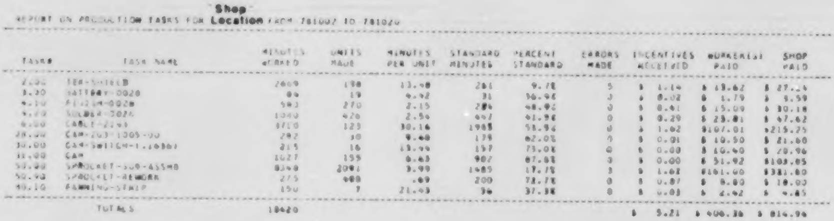
```
                    "Shop"
SHOP REPORT FOR Location FROM 781002 TO 781020
------------------------------------------------
TOTAL MINUTES WORKED                 60939
TOTAL INCENTIVES RECEIVED          $ 16.45
WORKERS PAID                     $  512.97
SHOP PAID                        $ 1028.19
SHOP PRODUCTIVITY                   32.48%
% TIME IN PRODUCTION                30.22%
% TIME IN PRODUCTION TRAINING        5.75%
% TIME IN SIMULATION                64.01%
% TIME IN SIMULATION TRAINING        0.00%
------------------------------------------------
```

FIGURE 2

Workshop Summary

minutes all workers spent in the production and training areas; the dollar amount the workers received as incentives under rate building individualized programs; the gross dollar amount the shop should have received for goods produced; the productivity of the workforce as a percentage of normal time standards for production tasks; and the percent of the total minutes worked that the workforce spent in production, training for production, simulation (unpaid time), and training for simulation (unpaid time).

Figure 3 is an example of a second kind of workshop report; the workshop (task x status) report. The example shown is a record of total workforce performance on production status tasks between October 2, 1978, and October 20, 1978. Similar reports (not shown) are produced for tasks classified as training for production simulation, and training for simulation. These reports give the task number and name; the minutes spent by the workforce on the task; the quantity of error free units they produced; the average number of minutes a workforce performing at one hundred per cent of the time standard would have required to produce the quantity actually produced; the average productivity of the workforce expressed as a percentage of the time standard for the task; the number of units produced that contained one or more errors; and three columns showing the incentives, workforce pay, and workshop income realized by the task.



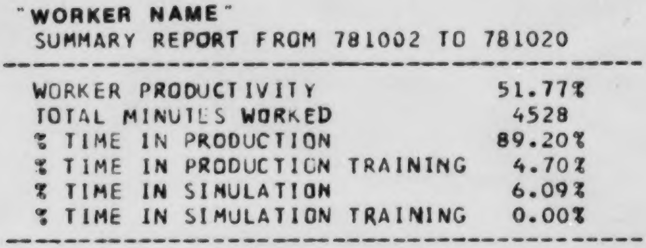
Shop

REPORT ON PRODUCTION TASKS FOR **Location** FROM 781002 TO 781020

FIGURE 3

Workshop (Task by status)

Worker Reports. These reports are similar to those produced for the workshop but contain information about an individual worker rather than the entire workforce. They are used by production supervisors, workshop directors, and agency personnel for setting and monitoring individual program objectives as well as for payroll and reimbursement activities. There are three kinds of worker reports; worker summary; worker (task x status); and worker by day. The first of these, the worker summary, is shown in Figure 4. This report contains the same information as that found in the workshop summary (Figure 2) with the omission of those items that refer to incentives, worker pay, or shop pay.

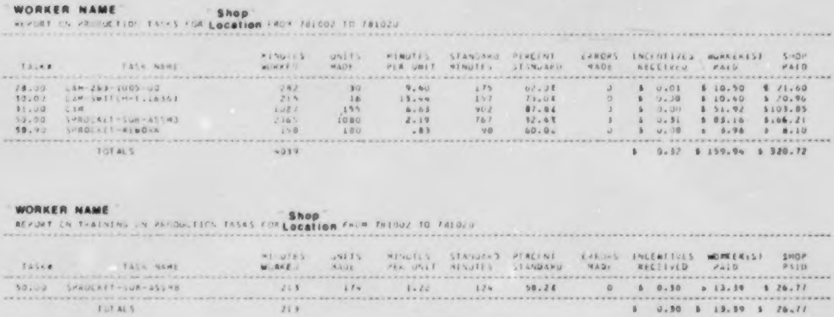


```
"WORKER NAME"
 SUMMARY REPORT FROM 781002 TO 781020
-----------------------------------------
 WORKER PRODUCTIVITY                51.77%
 TOTAL MINUTES WORKED                4528
 % TIME IN PRODUCTION               89.20%
 % TIME IN PRODUCTION TRAINING       4.70%
 % TIME IN SIMULATION                6.09%
 % TIME IN SIMULATION TRAINING       0.00%
-----------------------------------------
```

FIGURE 4

Worker Summary

The second kind of worker report, worker (task x status) is shown in Figure 5. Two examples of this report, one for "production" status tasks and the other for "training on production" status tasks are included in the figure.



"WORKER NAME"
REPORT ON PRODUCTION TASKS FOR **"Shop Location"** FROM 781002 TO 781020

TASK#	TASK NAME	MINUTES WORKED	UNITS MADE	MINUTES PER UNIT	STANDARD MINUTES	PERCENT STANDARD	ERRORS MADE	INCENTIVES RECEIVED	WORKER(S) PAID	SHOP PAID
78.00	CAM-263-1005-00	282	30	9.40	175	62.06	0	$ 0.01	$ 10.50	$ 21.60
50.01	CAM-SWITCH-1.16361	215	16	13.44	157	73.02	0	$ 0.00	$ 10.40	$ 20.96
51.00	CAM	1027	155	6.63	902	87.84	3	$ 0.00	$ 51.92	$103.85
50.00	SPROCKET-SUB-ASSMB	2365	1080	2.19	767	32.43	3	$ 0.31	$ 83.16	$166.21
58.90	SPROCKET-RIBBON	150	180	.83	90	60.00	0	$ 0.00	$ 3.96	$ 8.10
	TOTALS	4039						$ 0.32	$ 159.94	$ 320.72

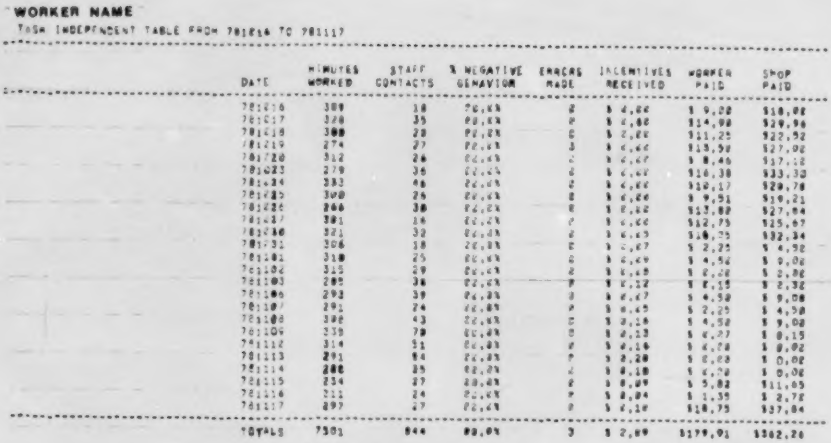
"WORKER NAME"
REPORT ON TRAINING ON PRODUCTION TASKS FOR **"Shop Location"** FROM 781002 TO 781020

TASK#	TASK NAME	MINUTES WORKED	UNITS MADE	MINUTES PER UNIT	STANDARD MINUTES	PERCENT STANDARD	ERRORS MADE	INCENTIVES RECEIVED	WORKER(S) PAID	SHOP PAID
50.00	SPROCKET-SUB-ASSMB	213	174	1.22	124	58.24	0	$ 0.30	$ 13.39	$ 26.77
	TOTALS	213						$ 0.30	$ 13.39	$ 26.77

FIGURE 5

Worker (Task x status)

An example of the third type of worker report, worker by day, is found in Figure 6. This report provides behavioral and earnings information for an individual worker on each day of the period covered by the report. For each day in the period the report gives the number of minutes worked on all tasks; the number of staff contacts the worker received; the proportion of observations in which target behavior was observed; the total number of defective units produced by the worker; the dollar amount earned from production and training work; and the expected workshop revenue resulting from that work.



```
WORKER NAME
TASK INDEPENDENT TABLE FROM 781016 TO 781117

         MINUTES   STAFF     % NEGATIVE  ERRORS  INCENTIVES  WORKER  SHOP
  DATE   WORKED    CONTACTS  BEHAVIOR    MADE    RECEIVED    PAID    PAID
```

FIGURE 6
Worker by Day

System summary

Both workshop and worker reports are generated from the STP offices. The reports required by a workshop, the time periods they cover, when they are to be generated, and the number of copies needed, are usually specified by workshop directors on a quarterly basis. However, the flexibility of the system allows additional reports to be produced at any time they are needed. Normal turnaround time from request to receipt of reports is four days (one day for generation, three days for mail service to the requesting workshop). Thus a workshop faces a total turnaround time from the day they mail input data to the day they receive output reports of approximately eight days.

The system has been operating for the workshops since May, 1978. Eight months of cost data for system operation show an average cost of eleven dollars per month per worker. This cost includes mailing, data entry, report generation, data storage, and software maintainance averaged across seventy-five workers in five workshops.

Future development

The STP information system in its current state is deficient in at least two respects. First, the turnaround time from data collection to report reception is too long, at least for information users such as production supervisors. Second, the system currently does not address the financial and demographic information required for effective workshop operation (examples of these sorts of information would be the source and amount of funding associated with individual workers, cost allocations, budget and income statements, inventory and production controls, etc.). In an attempt to resolve both of these deficiencies the next stage of system development will involve establishing a terminal at a selected site that will not only be able to use the STP programmatic report package described above, but also can access a specially tailored set of financial and demographic programs developed for the workshop. Such a terminal will constitute the first link in a distributed information system for STP workshops that will provide these small community based service settings with the information they must have if they are to provide effective service at a reasonable cost. With the continuing revolution in information processing (e.g., Branscomb, 1979) the prospects for such an undertaking look bright. To the extent these prospects can be confirmed empirically, we can look forward to a time when community service programs are linked together through computer technology into a system that benefits clients, service providers, and taxpayers alike.

References

Bellamy, G. T., Horner, R. H., & Inman, D. P. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Bellamy, G. T., Inman, D. P., and Horner, R. H. Design of vocational habilitation services for the severely retarded: The Specialized Training Program Model. In G. Hamerlynck (Ed.), Applied behaivor analysis techniques for the developmentally disabled. New York: Brunner Mazel, 1979.

Branscomb, L. M. Information: The ultimate frontier. Science 203, 1979, 143-147.

Schalock, R. L. The systems approach to program evaluation and accountability. Unpublished paper, 1978.

PART II

Stimulus Control
of
Vocational Behavior

Stimulus Control within Operant Chains: Variables Affecting the Intra-Chain Response Patterns of a Severely Retarded Woman

Robert H. Horner and G. Thomas Bellamy

The authors wish to express sincere appreciation to Barbara Wilcox for her review on earlier drafts of this article.

Preparation of this chapter was supported by the Bureau of Education for the Handicapped, Research Grant #G00-760-5455.

Stimulus Control within Operant Chains: Variables Affecting the Intra-Chain Response Patterns of a Severely Retarded Woman

Ordered sequences of behavior have long been of interest to researchers and theorists exploring functional descriptions of behavior (Kelleher, 1966; Millenson, 1967; Gollub, 1977). Research on behavior sequences has been particularly useful in the analysis of operant chains performed by severely retarded adults (Crosson, 1967; Horner & Bellamy, 1978; White, 1970). This research has focused on the role of intra-chain stimulus control. Within an operant chain, stimulus control describes the relationship between presentation of task-relevant stimuli and the probability that appropriate responses will be performed. Each step of the chain is viewed as a discriminative stimulus (S^D) → response (R) relationship. Responses within the chain do not occur until the particular S^D for a response is present. In addition, the performance of each response produces the S^D for the next response (Gollub, 1977). In this way the particular sequence of responses defining the chain is performed in a domino-like fashion until the chain is complete and a reinforcer is delivered.

Many educational and habilitation tasks require the performance of operant chains. Adding several numbers together, phonetically sounding out a word or assembling electrical components are but a few examples of operant chains. From an educational perspective the analysis of stimulus control within operant chains holds relevance for at least two reasons. First, the fast, accurate, independent performance of a task may only occur if careful attention has been given to developing stimulus control at each step in the chain. If the stimulus control within the chain is inadequate the habilitative value of training is substantially reduced.

Second, the likelihood of irrelevant behavior being performed within the chain is intimately connected to stimulus control. Any response which does not facilitate progress through the chain is defined as irrelevant. The performance of irrelevant responses is typically related to either inadequate stimulus control over appropriate responding, or the presence of non-task related stimuli which control irrelevant responses (Bellamy, Horner & Inman, 1979).

The present study provides a detailed single-subject analysis of stimulus control variables affecting the performance of relevant and irrelevant responses in a vocational chain performed by a severely retarded woman. The specific objectives of the study are to: (a) define those stimuli controlling a high-frequency, irrelevant response; (b) evaluate the presence of a functional relationship between increasing the level of reinforcement and the rates of appropriate and irrelevant responses; and (c) evaluate the presence of a functional relationship between a retraining strategy and the rates of appropriate and irrelevant responding.

Method

Subject

At the time of the study, the subject was 29 years old. She had been diagnosed as manifesting Downs Syndrome at birth and institutionalized since age 4. Institution records label her as severely mentally retarded, and administration of the Adaptive Behavior Scale (ABS) indexed her level of functioning at AAMD Adaptive Behavior Level IV. Approximately two years before the study began the subject was removed from her institutional residence and placed in a community group home facility with nine other severely retarded adults. At this same time she entered the Specialized Training Program, a contract workshop at the University of Oregon which provides training and employment to severely retarded adults (cf. Bellamy, Inman and Horner, 1979). The subject was completely non-verbal, but had been taught a small sign language (ASL) vocabulary, and was able to follow simple verbal directions.

Task

Approximately ten months prior to initiation of the study the subject was trained on a vocational task that involved pasting cushioning tape in the four inside corners of an 18 cm tall, rectangular, funnel-like apparatus used by Tektronix, Inc., to shield electronic elements from heat. Specifications for correct placement required that the tapes (a) measure no more than 4.4 cm and no less than 3.2 cm in length; (b) be positioned directly in the converging angles of the four corners; and (c) lie no more than .63 cm and no less than .3 cm below the outer lip of the heat shield. Completion of the task required correct performance of a 29 step operant chain. The individual discriminative stimuli and responses for each of these steps is provided in Table 1.

Industrial time standards suggest a non-handicapped worker should complete assembly of one heat shield every 1.2 minutes (i.e., at a rate of .83 per min.).

TABLE 1

Task Analysis of Heat Shield

STEP	S^D	RESPONSE
1	Shields and tape	Pick up shield
2	Shield in hand	Take off bag
3	Bag in hand	Throw bag away
4	Bag in garbage	Place shield on table
5	Shield on table	Pick up tape #1
6	Tape #1 in hand	Loosen tape backing
7	Backing loose	Throw backing away
8	Backing discarded	Pick up shield
9	Shield in hand	Put tape on corner
10	Tape in corner	Place shield on table
11	Shield on table	Pick up tape #2
12	Tape #2 in hand	Loosen backing
13	Backing loose	Throw backing away
14	Backing discarded	Pick up shield
15	Shield in hand	Put tape in corner
16	Tape in corner	Place shield on table
17	Shield on table	Pick up tape #3
18	Tape #3 in hand	Loosen backing
19	Backing loose	Throw backing away
20	Backing discarded	Pick up shield
21	Shield in hand	Put tape in open corner
22	Tape in corner	Place shield on table
23	Shield on table	Pick up tape #4
24	Tape #4 in hand	Loosen backing
25	Backing loose	Throw backing away
26	Backing discarded	Pick up shield
27	Shield in hand	Put tape in open corner
28	Tape in corner	Put shield in tray
29	Shield in tray	Raise hand

Irrelevant Responses

Preliminary observations indicated the subject performed several high frequency responses which were unrelated to task completion. Four of these: shield rocking, face touching, pausing and shrugging were identified as possible obstacles to the subject achieving a higher work rate. Operational definitions for each of these irrelevant responses are provided below:

Shield Rocking. Shield rocking was counted each time the subject tilted the shield forward or backward from its upright position on the work bench. Before a second instance of rocking could be counted the shield had to return to a stable, upright position.

Face Touching. Contact between the worker's hand and her face or hair. Before a second instance of face touching could be recorded the subject had to terminate contact with her face or hair.

Pausing. Three consecutive seconds during which the worker did not exhibit any work-related responses or perform another targeted irrelevant response.

Shrugging. Any movement in which the worker either moved both shoulders forward while simultaneously pushing her head back; or moved both shoulders back while simultaneously pushing her head forward.

Design

A single subject design incorporating multiple reversals (Hersen & Barlow, 1976; Kratochwill, 1978), was used to evaluate stimulus control of relevant and irrelevant responses throughout the 29 step task. The dependent variables were (a) the rate at which shields were completed, (b) the rate of each of the four irrelevant responses, and (c) the conditional probability of each of the irrelevant responses at each step in the operant chain. The independent variables were (a) type and schedule of reinforcement; (b) a retraining procedure to reduce irrelevant responding, and (c) presence of specific supervisors.

Setting

The study was conducted in the production room of the Specialized Training Program. Within this environment fourteen severely handicapped workers assembled electronic components at individual work stations. Workers were in production during three work periods daily. Each work period was approximately one and one-half hours long and ended with a break. During the work periods two supervisors circulated throughout the workshop to (a) ensure that workers had adequate work materials, (b) check work quality, and (c) deliver social praise and tokens. Tokens were delivered contingent upon accurate work completion and social praise was delivered contingent upon on-task behavior. In general, the workshop was a highly contingent environment in which reinforcers were delivered only after appropriate behavior, and potential reinforcement of inappropriate behavior was minimized.

Observation

The subject performed the heat shield task throughout her work day. Data from two trials were collected each work day morning by observing the worker from behind a one-way mirror located immediately above and beside her. A trial consisted of completing one shield. Data collected during a trial included the time required to complete the shield and the frequency of each irrelevant response at each step in the operant chain. When the subject reached for the shield a stopwatch was started. When she performed the last step of the operant chain (i.e., raised her hand signaling completion) the stopwatch was stopped. The number of minutes required for completion was then entered into the formula, $\frac{1}{\text{Min. to completion}}$, to obtain a measure of work rate.

As the subject performed the task, a frequency count was also kept for each irrelevant response at each step in the chain. An irrelevant response was recorded as occurring during a particular step if it was performed after the S^D for that step was present, but before the S^D for the next step was present. Irrelevant responses other than those defined above were not recorded. The conditional probability of each irrelevant response at each step in the chain was calculated by dividing the number of times the response occurred at least once during a step by the number of trials in the experimental phase. This method of calculating conditional probability indexed the likelihood of each irrelevant response at each step in the chain.

Reliability

Thirty-eight times during the study, data were simultaneously collected by two observers sitting at separate observation windows above the workshop. Reliability on the rate of shield completion was determined by computing percent agreement across all reliability sessions over the study. If the times recorded by the two observers for shield completion were within 1 sec of each other they were scored as an agreement. Percent agreement was calculated by taking agreements divided by agreements plus disagreements, and multiplying by 100.

Reliability on the rate of irrelevant responses was computed by percent agreement for each irrelevant response. The frequencies recorded by two observers for each irrelevant response at each step in the operant chain were compared. If the same number of responses had been recorded by both observers for a step an agreement was scored. If the frequencies differed a disagreement was scored.

Four graduate students in Special Education served as observers. One observer conducted all reliability checks. No observer's data was included in the study until s/he had demonstrated 90% reliability over two consecutive days on all dependent measures.

Procedures

The present study is a longitudinal analysis of variables functionally related to one severely retarded woman's production rate. It is comprised of ten experimental phases. Phases 1-4 form an ABAB reversal design examining the effects of increasing reinforcement for completion on the rate of appropriate and irrelevant responses. Phase 5 examines changes in the same dependent variables as a function of specific supervisors being present. The final five phases (6-10) focus on the utility of one easily implemented intervention for reducing irrelevant responding. Procedures for each of the ten phases are described below:

1. **Baseline 1 (BL 1).** During this phase the subject received social praise and a penny for each shield correctly completed. Shields that were incorrectly assembled were returned to the subject for correction. Pennies were exchanged for nuts, raisins or small candies at the end of the first and third work periods.

The second period ended at noon and was followed by lunch. The BL_1 phase lasted six days during which eleven trials were observed.

2. **Immediate Exchange 1 (IE 1).** Throughout the IE_1 phase the subject was able to immediately exchange her pennies for a variety of liquids (coffee, tea, orange juice and Coca Cola) that were identified by staff in her living environment as high preference consumables. Approximately 1 oz. of a liquid was exchanged for a penny immediately after the penny was earned. The liquids were frequently varied in an attempt to minimize satiation. This phase lasted six days, during which eleven trials were observed.

3. **Baseline 2 (BL 2).** Procedures for this phase replicated the procedures for BL_1. These procedures were in effect four days during which seven trials were monitored.

4. **Immediate Exchange 2 (IE 2).** This phase replicated IE_1. Two days (4 trials) into this phase, however, Supervisor A left for vacation and was out of the production environment during the remainder of the IE_2 phase. IE_2 lasted thirteen days during which twenty-five trials were monitored.

5. **Supervisor Preference (SP).** Erratic data following Supervisor A's departure during the IE_2 phase suggested the possibility that the subject was responding differentially to production supervisors. The SP phase examined this possibility. A staff trainer who had not been on the production floor during earlier phases was asked to work as a production supervisor (Supervisor B) every other day during the first 7 days (15 trials) of this phase. This supervision schedule served as a reversal design to assess the effects of Supervisor B's presence on the subject's work behavior. Supervisor B had been a production supervisor approximately six weeks before the study began, and had worked as a trainer with the subject on an earlier task. He did not serve as a supervisor beyond the initial 15 trials of this phase.

Fifteen trials into the SP phase Supervisor A returned from vacation and resumed his role as a production supervisor. To evaluate the effect of Supervisor A's presence on the subject's performance the final 25 trials of the SP phase were manipulated in a BABAB reversal design with Supervisor A's presence in the production room serving as the independent variable. Throughout the 40 trials (20 days) of the Supervisor Preference phase Supervisors A and B were aware that the subject's response to their presence was being monitored. To minimize the effect this might have on their performance, both supervisors were instructed not to alter the topography of their interactions with the subject from previous patterns.

6. **Retraining.** During the Retraining phase an attempt was made to alter the stimulus control that task-relevant stimuli were exerting over the subject's shield rocking response. Supervisor C, who had been working as a production supervisor across all previous phases, implemented the intervention. The supervisor stood immediately behind the subject, and each

time the shield was rocked she would immediately remove it, say "No rocking" and walk away. Approximately 30 sec later the shield was returned to the subject and Supervisor C resumed her monitoring position. The subject was not consequated for rocking by any other supervisor at any point during this phase. Supervisor C did not enter the production room except in the second morning work period to conduct retraining sessions. Data from this phase were only collected while retraining was in progress. The time recorded for shield completion includes the time consumed while rocking interventions were carried out.

7. **Supervisor C.** During this phase the subject's differential responding to Supervisor C was assessed. Supervisor C did not stand behind the subject and, except for the first six trials of the phase, did not consequate rocking. During the first six trials Supervisor C carried out the same intervention procedures used during Retraining. Throughout this phase Supervisor C was instructed to function as a "normal" supervisor. She was instructed not to constantly monitor the subject, or make a special effort to "catch" all instances of the rocking response. All other supervisors continued to ignore shield rocking.

Following trial six of this phase, Supervisor C ignored any rocking. Her contact with the subject from this point on did not differ from those of other supervisors. The presence of Supervisor C was manipulated throughout the phase in a reversal design to assess any functional relationship between the subject's work performance and Supervisor C's presence. The Supervisor C phase lasted 29 trials (15 days). Supervisor C did not work as a supervisor following the conclusion of this phase.

8. **Staff Consequation of Rocking (SCR).** Procedures for this phase were in effect across 49 trials (25 days). This phase was designed to assess the degree to which production rate gains observed in the two previous phases would maintain with all staff implementing the intervention used by Supervisor C.

During the initial 41 trials of this phase (Trials 132-174) each work period started with a supervisor standing behind the subject as she completed one shield. The subject was praised if she did not rock the shield, and if she did emit a rocking response the shield was taken from her, the supervisor said "No rocking," walked away for approximately 30 sec, and then returned the shield. Care was taken to ensure that all supervisors had an equal opportunity to stand behind the worker as she completed the first shield of a work period.

After the first shield had been completed the supervisor returned to his or her normal supervision duties. If, following this one trial, the subject was observed rocking the shield the supervisor observing the rocking would implement the intervention procedure. As in the previous phase the frequency of staff consequation for rocking was monitored via a tally sheet at the subject's work station. Following Trial 174, supervisors did not stand behind the subject at any time. They did, however, continue to consequate instances of rocking.

9. **Supervisor D**. During this phase a fourth supervisor (Supervisor D), who had been working on the production floor throughout the study, was designated as the only supervisor who would intervene when the subject paused. The same intervention used in the previous three phases for rocking responses was used during this phase for pausing. If the subject did not engage in work-related responses or other targeted irrelevant responses for three consecutive seconds, Supervisor D would take away her shield, say "You are not working," walk away for approximately 30 sec and then return the shield to her. Supervisor D functioned in a normal supervisory role throughout the phase. He was instructed not to stand behind the subject or increase the overall frequency of his attention to her. To assess the effect this alteration in supervisor behavior might have on the subject's behavior an ABABABABAB reversal design was conducted during the phase, with presence of Supervisor D serving as the independent variable. Supervisor D used a tally sheet to monitor the number of times per day that he consequated the worker for pausing. This phase lasted 24 trials (12 days).

10. **Staff Consequation of Pausing (SCP)**. During the final phase of the study all supervisors were instructed to follow the intervention procedures used by Supervisor D if they observed the subject pause. This phase lasted 37 trials (20 days).

Results and Discussion

The reliability of data for production rate and the rate of each irrelevant response was assessed 38 times during the study. As shown in Figure 1, observer agreement scores were collected in all phases. Observers recording the number of minutes to complete a shield were within one second of each other 37 of the 38 observations, for an agreement score of 97%. The one discrepancy was a three second difference.

Observer agreement on the frequency and position in the operant chain for each of the irrelevant responses (shield rocking, pausing, face touching and shrugging) ranged between 89% and 100%. The average level of agreement for each irrelevant response in each phase of the study was over 92%.

Section One: Stimulus Control of Irrelevant Responding under Increased Reinforcement for Work Completion

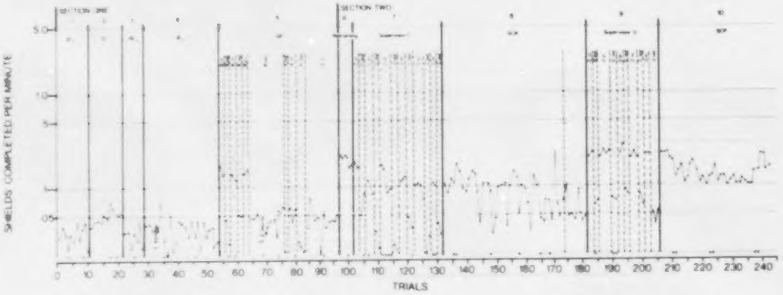
The first four phases of the study provide an ABAB design examining change in production rate and irrelevant responding as a function of increasing the level of reinforcement for work completion. Results indicate that increased reinforcement is related to improved rates of production, though the effect appeared to interact with the presence of specific supervisors. To further examine this interaction a fifth phase (the SP phase) was added to the design. Results from all five phases of section one will be reported first for production rate, and then for irrelevant response rates.

Production Rate. As shown in Figure 1, the subject produced shields at a stable rate of approximately .03 shields per minute (i.e., 33 min per shield) across the 11 trials of BL_1. During the IE_1 phase both the type of reinforcement and schedule of exchange were varied in an attempt to maximize the likelihood that the Immediate Exchange phase would provide an increased level of reinforcement for work completion behavior. If this objective were successful an increase in production rate would be expected when immediate exchange procedures were implemented. This is, in fact, the performance pattern observed. The trend of the data for IE_1 demonstrates a sharp increase in production to a level nearly double that of the BL_1 phase. This represents a reduction of nearly sixteen minutes in the time required to complete each shield.

During the BL_2 phase procedures returned to those used in BL_1. Immediately following this change the subject's rate of production dropped back to the BL_1 level, and remained at an average of .03 shields per min throughout the seven trials of the BL_2 phase. This decrease in production rate represents a dramatic change in both trend and level from that observed during IE_1.

The subject's production rate data across the BL_1, IE_1, BL_2 phases follows an ABA reversal pattern. The immediacy of rate changes following phase changes and the magnitude of rate improvement in IE_1 argue strongly for a functional relationship between IE_1 procedures and an increase in production rate. Given that both type of reinforcement and exchange schedule were manipulated simultaneously, however, there is no experimental control to assess the individual effects of these two variables.

When immediate exchange procedures were reintroduced in the IE_2, it was expected that rate improvements similar to those observed in IE_1 would occur. The first four trials of IE_2 were, in fact, slight improvements over the baseline levels in the BL_2 phase. Four trials into IE_2, however, Supervisor A who had been the major supervisor in the workshop during previous phases left on vacation. He spent more time per day on the production



Triangles on the abcissa indicate trials when reliability was taken across all phases.

FIGURE 1

Rate of Shields Completed per Trial

floor than did other supervisors and had been involved in previous one-to-one training with the subject. Following Supervisor A's departure, the subject's rate of production initially increased and then dropped to a level similar to that observed in the previous baseline phases (.034 shields per min.).

The marked improvement in rate noted in IE_1 was not replicated in IE_2. This suggests the possibility that (a) the effect noted across the BL_1, IE_1, BL_2 phases was not functionally related to experimental procedures or (b) changes in uncontrolled variables during the IE_2 phase inhibited rate improvement. The fact that within 2 trials (1 day) after Supervisor A's departure the subject's rate both lowered in level and increased in variability suggests that the presence of Supervisor A may have been one stimulus controlling higher performance rates.

To examine this hypothesis the Supervisor Preference (SP) phase (phase 5) was conducted. The SP phase data presented in Figure 1 indicate that the subject's rate was affected by the presence of both Supervisors A and B. The multiple reversals in the initial trials of the SP phase suggest a strong functional relationship between improved work rate and Supervisor B's presence. These data suggest that the subject's production performance was affected by the presence of individual supervisors, thereby supporting the hypothesis that the IE_2 data may have been affected by Supervisor A's departure.

Fifteen trials into the SP phase Supervisor A returned, and resumed his role in the production room. With Supervisor A's return the SP procedures exactly replicated those of the earlier immediate exchange phases. During the six trials immediately following Supervisor A's return the data indicate a gradually improving trend in production rate. The improved performance across these trials closely resembled the change in trend and level noted in IE_1. Her rate during trials 73, 74 and 75 was approximately .06 shields per min. This is a substantial increase over the .03 shields per min rate observed in the Baseline phases. To assess the degree to which this improvement was functionally related to the presence of a particular supervisor a BABAB design was implemented with Supervisor A's presence manipulated as the independent variable. The data from these reversals indicate immediate and substantial shifts in production rate following each experimental manipulation. Production rate maintained at .05 to .06 shields per min in Supervisor A's presence and dropped to the extremely low rate of approximately .015 shields per min in his absence. These results indicate that Supervisor A's presence on the production floor was functionally related to variability in the subject's work rate.

In general, production rate data across the BL_1, IE_1, BL_2, IE_2 and SP phases argues that both the presence of Supervisor A and the Immediate Exchange procedures affected the subject's performance. With Supervisor A's presence held constant implementation of IE procedures within an ABA reversal design (BL_1, IE_1, BL_2) indicated a strong functional relationship between IE_1 procedures and improved production rate.

Of special interest with respect to the subject's differential performance to supervisors was her response to Supervisor B. Results from the BABAB reversals in the early trials of the SP phase indicate immediate improvement in production rate when Supervisor B was present. Production rate without Supervisor B maintained at approximately .03 shields per min, and increased to well over .1 shields per min with Supervisor B present. This not only lends support to the hypothesis that the subject was responding differentially to specific supervisors, but provides a clear index of the subject's ability to produce at rates substantially above her baseline level.

Data from the first five phases indicate two independent variables related to production rate (reinforcement of task completion behaviors and the presence of specific people in the production setting). Both of these findings have been documented extensively elsewhere (Bellamy, Inman & Yeates, 1978; Evans & Spradlin, 1966; Horner, 1977; Huddle, 1967; Redd, 1976; Redd & Birnbrauer, 1969; Risley, 1968). Little attention, however, has been directed at the effect of these independent variables on the performance of irrelevant behaviors. The present study provides a fine-grained index of their effect on the rate and probability of four irrelevant responses.

Irrelevant Response Rates. Data on the rate of rocking, pausing, face touching, and shrugging across the ten phases of the study are presented in Figures 2-5. These data are discussed first with respect to the effect of increasing reinforcement for task completion (section one), and then with respect to differential responding to supervisors (section two).

During the 11 trials of BL$_1$ rocking, pausing, face touching and shrugging all exhibited stable rates of approximately 4.5 rockings per minute; well over 100 instances of rocking per shield. The rate of pausing during BL$_1$ was initially more variable than rocking. The last five trials of the BL$_1$ phase, however, evidence a stable pausing rate of approximately .46 pauses per min.

The rate of face touching and shrugging followed a variability pattern similar to pausing. The initial trials of the BL$_1$ phase indicated increases in rate and variability, while rates during the last five to seven trials were stable for both irrelevant responses (.55 instances per min for face touching and .72 responses per min for shrugging).

It was anticipated that these stable rates would follow an ABA pattern across the BL$_1$, IE$_1$, BL$_2$ phases. As production rate increased it was expected that the rate of irrelevant responses would decrease. This pattern proved predictive of shrugging, face touching, and to a lesser extent pausing. As seen in Figures 4 and 5 , shrugging and face touching responses evidenced downward trends during IE$_1$. Both irrelevant responses were performed at less than half their baseline levels at the end of IE$_1$. With the introduction of BL$_2$ and the concomitant reduction in production rate the rate of both face touching and shrugging immediately increased. Across BL$_2$ the rate of face touching averaged .94 instances per min, more than three times the rate observed during BL_1. The rate of shrugging during BL_2 also follows a reversal in trend from the IE_1 phase. As can be seen in Figure 5, the subject's rate of shrugging gradually returned to BL_1 levels over the 7 trials of BL_2.

Together the shrugging and face touching results conform closely to the pattern expected. As the subject performed her vocational responses faster, the rate of irrelevant shrugging and face touching decreased. When production rate dropped back to baseline levels the rate of shrugging and face touching increased.

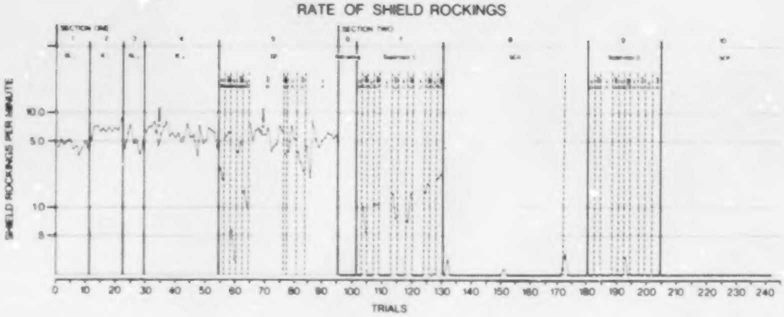


FIGURE 2

Rate of Shield Rockings per Trial

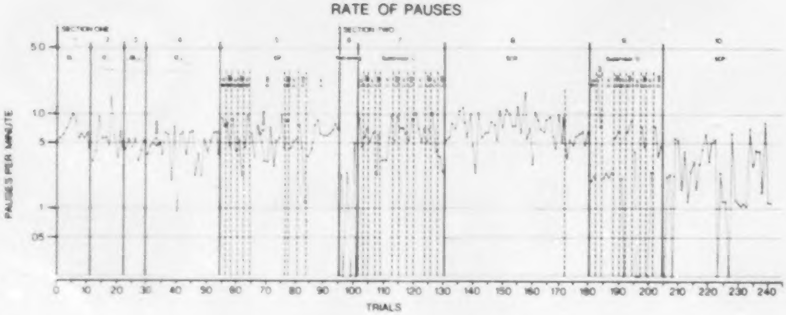


FIGURE 3

Rate of Pauses per Trial

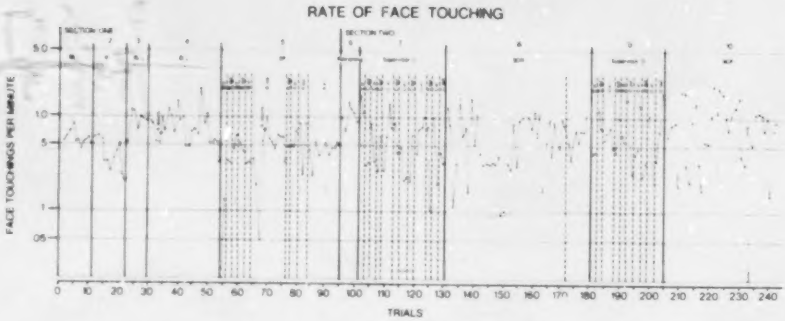


FIGURE 4

Rate of Face Touching per Trial

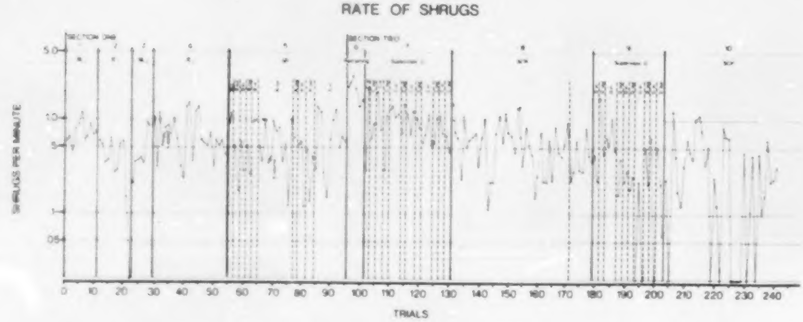


FIGURE 5

Rate of Shrugging per Trial

There was no evidence of a functional relationship between pausing and reinforcement manipulation across BL$_1$, IE$_1$, and BL$_2$. During these phases the subject averaged .67, .59, and .46 pauses per min, respectively. While the rate of pausing was not affected by the experimental manipulations, observers did report that the duration of pauses during the IE$_1$ phase were generally shorter than during the baseline phases. However, since data were not collected on pause duration verification of these reports is not possible.

As can be seen in Figure 2, rocking followed yet a third response pattern across the first four phases of the study. Increasing reinforcement for task completion was functionally related to an increase in the rate of rocking. A mean BL$_1$ rate of 4.5 rocking responses per min rapidly increased to an average rate of 6.7 responses per min during IE$_1$. With the return to baseline in BL$_2$ rocking dropped in level, and changed in trend such that within 7 trials the rocking had returned to BL_1 levels. During the first 4 trials of IE_2 (i.e., before Supervisor A left the workshop), there was an immediate and sharp increase in the rate of rocking. This data argues strongly for a functional relationship between an increased rate of rocking and reinforcement for work completion. This is somewhat puzzling given that rocking in no way facilitated rapid task completion. In fact, the 100+ rocking responses performed per shield added substantially to the time required to complete each shield.

The Conditional Probabilities of Irrelevant Responses. The different performance patterns of the irrelevant responses across IE, BL and IE_2 become more clear when viewed in conjunction with the conditional probability data. Figure 6 presents the probability of rocking, pausing, face touching and shrugging at each step in the operant chain. While these data are from the BL_1 phase, they are representative of the conditional probability data across all phases. Comparisons indicate that while pausing, face touching and shrugging were highly probable throughout the chain, shield rocking occurred only at certain steps. At these steps, however, the probability of rocking following presentation of the S^D for the step was very high.

Further analysis of this pattern indicates that of the five steps for which rocking was 100% probable, all were preceded by the response of placing the shield down. That the subject rocked the shield each time she set it down, and seldom rocked it any other time suggests that rocking was, in fact, under stimulus control of seeing and/or feeling the shield placed on the work bench. If these task-relevant stimuli did in fact control rocking, it would be expected that the rate would increase as the rate of production increases, since increasing production rate would simply increase the frequency with which the S^D for rocking was presented.

The combined analysis of the rate and probability of irrelevant responding provides a much finer perspective of rate variability across the BL_1, IE_1 and BL_2. Clearly, increased reinforcement for task completion was effective in increasing production and in maintaining or decreasing the rate of irrelevant behaviors not under control of task stimuli. For the one response that was controlled by stimuli inherent to the task, increasing the level of reinforcement for task completion was related to an increase in the rate at which the response was performed.

The performance of irrelevant responses across the IE_2 and SP phases provides an index of the effect of supervisor presence. The rate of rocking during the IE_2 phase was well above baseline levels, and considerably more variable than any of the previous phases. The first four trials of IE_2 indicated a marked increase in rocking over the BL_2 rate. After Supervisor A departed, however, the subject's rate of rocking dropped and then followed a highly variable pattern, averaging 5.9 rocking responses per min over the remaining 21 trials. This is above the 4.6 per min rate observed during baseline, yet well below the mean 6.7 response per min performed during IE_1.

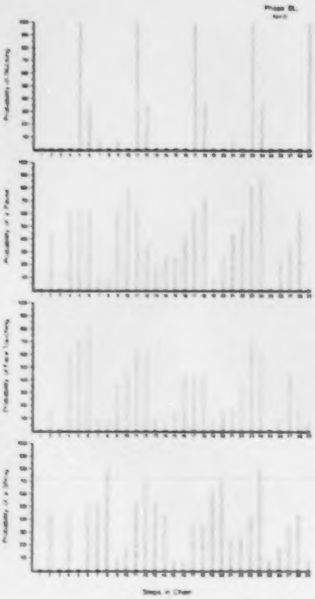


FIGURE 6

The Conditional Probability of each Irrelevant Response of each Step in the Operant Chain

Rocking data for the SP phase indicate that Supervisor B exhibited substantial control over the rate of rocking while Supervisor A did not. During the initial reversals with Supervisor B, the subject's rate of rocking dropped to one-fifth its baseline rate when Supervisor B was present and immediately returned to baseline level when Supervisor B was absent. As noted in Figures 2 and 3, these reductions in rocking rate were directly correlated with the highest levels of production during phases 1-5. Reversals later in the SP phase which manipulated the presence of Supervisor B indicate consistently improved performance in Supervisor B's presence, though the short duration of each phase, and the minor shifts in level across phases suggest that this control was not significant.

Results for pausing, face touching, and shrugging over the IE_2 and SP phases are presented in Figures 3, 4 and 5. The rate of each of these irrelevant responses is extremely variable over the IE_2 and SP phases. No functional relationship between rate of irrelevant responding and supervisor presence is apparent.

Section Two: Supervisor Control of Irrelevant Responding

Section two results are taken from the last five phases of the study and constitute an examination of supervisor control over irrelevant responding. Retraining, Supervisor C, and SCR phases all examine the effectiveness of a procedure for reducing the stimulus control of task-relevant stimuli over rocking and its impact on production rate. The Supervisor D and SCP phases used the same intervention in an attempt to decrease the subject's rate of pausing, and the impact of this manipulation on production rate.

Stimulus Control of Rocking. During Retraining, the subject was not allowed to continue through the chain if she rocked the shield in the presence of Supervisor C. If she rocked the shield, Supervisor C removed it thereby terminating the subject's access to intra-chain task-relevant stimuli for approximately 30 sec. In the presence of Supervisor C the subject received access to liquids only by performing the operant chain without irrelevant rocking.

Results from Retraining indicate an immediate change in the subject's production behavior from SP levels. Production rate increased sharply, as did the rate of face touching and shrugging, while rocking and pausing decreased.

Production rate data provided in Figure 1 show a level increase in the subject's rate of production of 400% above that observed in the preceding 12 trials of the SP phase. Concomitantly, the rate of rocking dropped from approximately 5 per min during SP to a consistent 0 responses per min during Retraining.

Rate of pausing, which had been very stable across all previous phases, decreased substantially when Retraining procedures were implemented. From a level of .58 pauses per min during the SP phase the subject's performance dropped to a mean of .16 pauses per min during Retraining. In addition to this decrease there was a marked increase in the variability of pausing during Retraining.

The rate of face touching and shrugging both doubled when Retraining procedures were implemented. Face touching occurred at a mean rate of 1.17 responses per min while shrugging occurred at a rate of 2.2 responses per min. These rate increases are opposite the effect noted across the BL_1, IE_1, BL_2 phases when production rate increased, and the rate of face touching and shrugging decreased. These levels were not maintained in the Supervisor C phase, however, making it unlikely that the interruption procedure was the sole variable related to the increased rates of face touching and shrugging. The present design does not afford an analysis of other Retraining variables that may have affected face touching and shrugging.

Following the implementation of Retraining procedures, an AB effect was observed across all dependent variables. To determine if Retraining procedures were functionally related to performance change when used in a normal supervisory context the Supervisor C phase was implemented. During this phase Supervisor C followed normal supervision patterns. Across the first three days of the phase, however, she continued to intervene when the subject rocked the shield. Supervisor C consequated rocking only four times during these three days and with none of these instances occurring during data collection trials. At no point beyond this time was the subject observed to exhibit rocking in the presence of Supervisor C though rocking did occur when Supervisor C was not present.

Results from the multiple reversals of this phase are presented in Figures 2 and 3 and clearly indicate that Supervisor C's presence had a major impact on both the subject's production and rocking rates. When Supervisor C was present the subject produced an average of .12 shields per min with no instances of rocking. In Supervisor C's absence the subject's rate dropped to a mean level of .035 per min and included an average of 1.33 instances of rocking per min. The immediate level changes across the multiple reversals in the Supervisor C phase provide strong support for a functional relationship between Supervisor C's presence, increased production rate and a decrease in rocking rate. Significantly, the subject maintained this strong reversal pattern across the latter three-quarters of the Supervisor C phase even though Supervisor C's behavior did not deviate in observable ways from that of other supervisors. Prior to Retraining the subject had not responded differentially to Supervisor C. Unlike the functional relationship between production rate and rocking no differential effects were noted for pausing, face touching or shrugging across the reversals in the Supervisor C phase. Each of these responses occurred at rates approximating those exhibited prior to Retraining.

The results from multiple reversals of the Supervisor C phase indicate that the interruption of rocking behavior was sufficient to reduce the rate of rocking. This suggests the possibility that, as a result of interrupting the chain, Supervisor C came to function as a stimulus with the capacity to decrease the strength of the rocking response. In the presence of Supervisor C, the control that task-stimuli previously held for rocking was effectively inhibited.

The performance of rocking in Supervisor C's absence demonstrates that control of rocking by task-relevant stimuli still existed during this phase. As seen in Figure 2 data on the rate of rocking during Supervisor C's absence demonstrates a gradual upward trend across the phase. This would be consistent with the subject's "learning" that only in the presence of Supervisor C would rocking be interrupted.

Results from the Supervisor C phase suggest that if all supervisors began interrupting the chain when rocking occurred there should be a general decline in the rate of rocking. SCR phase in which all supervisory staff consequated shield rocking examined this hypothesis. During the initial 21 trials of SCR production rate stabilized at .119 shields per min. The remaining trials indicate a gradual decline in rate to a mean level of .089 shields per min which, nonetheless, represents a three-fold increase over baseline production rate levels.

It is apparent from Figure 2 that the rate of rocking during SCR was negligible. Subjective reports from supervision staff substantiate these data. Given the correspondence between rocking and production rate, it is likely that the subject's improved production rate during SCR was directly related to the low level of irrelevant rocking.

Staff intervention for rocking occurred four times during the first day of the SCR phase, twice on each of the next two days, and a total of five times during the remaining 19 days of the phase. Rocking was not observed during the last 8 days of SCR. These data suggest that the previous control that task stimuli had emitted over rocking responses was either substantially reduced, or inhibited. Conversely, results shown in Figures 3, 4 and 5 indicate that SCR procedures did not have any effect on the rate of pausing, face touching or shrugging.

Viewed together, data from Retraining, Supervisor C and SCR indicate that: (a) as a function of supervisor interruption of rocking, the subject's rate of rocking decreased to a virtually non-existent level; (b) reduction in the rate of rocking was directly correlated with substantial increases in production rate; (c) the subject very rapidly learned to respond differentially in the presence of a supervisor who had not been correlated with differential responding previously; and (d) staff performing their normal supervisory duties were able to implement the intervention procedures effectively.

Stimulus Control of Pausing. In the opinion of observers, the subject's slight decline in production rate during the SCR phase was related to an increase in the duration of pauses. The Supervisor D phase represents a response to this input. The intervention procedure that had effectively reduced the rate of rocking during previous phases was applied to pausing. To determine if this intervention was functionally related to change in the rate of production and irrelevant responding, the interruption procedure was implemented in a multiple reversal design with the presence of Supervisor D (who implemented the procedure) serving as the independent variable.

As can be seen in Figure 1 production rate across the reversals remained at the SCR level of .089 shields per min when Supervisor D was not present. When Supervisor D was in the production room interrupting pauses, however, the subject's production rate averaged a stable .226 shields per min. This represents the highest, and most stable, rate of performance the subject had produced up to that point. Both the immediacy and magnitude in level changes across Supervisor D reversals argues convincingly that Supervisor D's interruption of pausing was functionally related to improved production performance.

Results provided in Figure 3 suggest that this improved level of rate with Supervisor D present may be a function of reduced rates of pausing. Across the multiple reversals in this phase the subject performed pauses at a much reduced rate when Supervisor D was in the production room. Observers also indicated that the duration of pauses seemed much shorter when Supervisor D was present.

Supervisor D consequated pausing 8 times on the first day of the Supervisor D phase, 5 times on each of the next two days and 2 or 3 times each day for the remainder of the phase. Unlike the Supervisor C phase there was no time at which the subject's rate of pausing dropped so low that the interruption procedure could be discontinued.

During the SCP all staff interrupted the subject when they observed her pause. Throughout SCP production staff consequated the subject for pausing at least once a day. The mean level of consequations per day for the entire phase was 2.6, with a range of one to six. Production rate during this phase declined somewhat from the high rates of the previous phase to an average of .139 shields per min. However, this level of performance was four times higher than that observed during the initial baseline phases, was nearly double the subject's production rate during the SCR phases, and maintained for over four weeks.

Rate of pausing during the SCP phase was extremely variable ranging from 0 instances per min to .8 instances per min. In addition, the overall rate of pausing for the SCP phase was less than that observed for any other phase.

Across both Supervisor D and SCP phases, the rate of rocking remained at 0. No functional relationships were evident between independent variable manipulations and the rate of either face touching or shrugging during these phases. There was, however, a general decrease in level, and increase in variability for the rate of shrugging during SCP.

Results from the Supervisor D and SCP phases indicate that interruption of pausing was effective in both reducing the subject's rate of pausing and in increasing production rate. This effect was much less dramatic, however, than that observed when the same intervention was applied to rocking, and at no point were staff able to maintain a high production level without continually intervening on pausing.

The strategy of interrupting irrelevant responses could have been less effective with pausing for several reasons. First, a retraining phase was not conducted with pausing, thereby limiting the initial precision with which procedures were implemented. A related explanation is that pausing may be more difficult for supervisors to observe than rocking, with the result that more instances of pausing passed unnoticed. If pausing were not under control of task-relevant stimuli, and, therefore, not part of the operant chain, the interruption of pauses may not have been as punishing as interruption of rocking, or as effective in reducing the control exerted by work related stimuli. A final hypothesis is that pausing is a response that occurred in many settings during many tasks. If this were the case the subject may have been reinforced for pausing at other times of the day. It is very unlikely, however, that shield rocking occurred except in the context of the production environment.

Summary

Of major importance in the present study is the effect of irrelevant responding on the subject's production rate. In general, the higher the rate of irrelevant responding the lower the subject's rate of production. Further, one irrelevant response (rocking) was actually under control of task-relevant stimuli. The subject behaved, in effect, as if performance of rocking was actually part of the operant chain defining the task. Manipulation of reinforcement level did not reduce the dilatorious impact of rocking on the subject's rate of production, though it did reduce the rate of those irrelevant responses not controlled by task-stimuli. An easily implemented interruption procedure effectively reduced the control of task-stimuli over rocking, resulting in an immediate and substantial improvement in production rate. The same procedure used with pausing was also effective in improving production rate and decreasing the subject's rate of pausing. These data suggest that important improvements in the work rate of some severely retarded workers will require more than gross manipulation of reinforcement for work completion. Of central concern in programming for retarded workers who work slowly must be an assessment of stimulus control within the operant chain defining their task.

An emphasis on evaluating a worker's performance of operant chains shifts the focus of habilitation analysis. The traditional approach of treating the entire chain as a unit of responding (i.e., by only assessing global rate measures) is certainly necessary but not sufficient for a complete appraisal of important variables affecting the vocational training and supervision of severely retarded people. A more complete assessment will require analysis of worker intra-chain performance patterns.

References

Bellamy, G. T., Horner, R. H., and Inman, D. P. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Bellamy, G. T., Inman, D. P. and Horner, R. H. Design of vocational habilitation services for the severely retarded: The Specialized Training Program Model. In G. Hamerlynck (Ed.), Applied behavior analysis techniques for the developmentally disabled. New York: Brunner Mazel, 1979.

Bellamy, G. T., Inman, D. P. and Yeates, J. Workshop supervision: Evaluation of a procedure for production management with the severely retarded. Mental Retardation, 1978, 16, 317-319.

Crosson, J. The experimental analysis of vocational behavior in severely retarded males. Final report, Grant No. OEG 32-47-0230-6024, Washington, D. C.: Department of Health, Education and Welfare, 1967.

Evans, G., and Spradlin, J. Incentive and instructions as controlling variables in productivity. American Journal of Mental Deficiency, 1966, 71, 129-132.

Gollub, L. Conditioned reinforcement: Schedule effects. In W. K. Honing and E. R. Staddon (Eds.), Handbook of operant

behavior. New Jersey: Prentice-Hall, 1977, p. 288-309.
Hersen, M. and Barlow, D. Single case experimental designs. New York: Pergammon Press, 1976.
Horner, R. H. Stimulus control, transfer and maintenance of upright walking posture with a severely retarded adult. In G. T. Bellamy, R. H. Horner, and D. P. Inman (Eds.), Habilitation of severely and profoundly retarded adults: Volume II. Eugene, Oregon: University of Oregon, Center on Human Development, 1977.
Horner, R. H. and Bellamy, G. T. Structured Employment: Productivity and productive capacity. In G. T. Bellamy, G. O. O'Connor, and O. C. Karan (Eds.), Vocational habilitation for developmentally disabled persons: Contemporary service strategies. Baltimore: University Park Press, 1978.
Huddle, D. Work performance of trainable adults as influenced by competition, cooperation, and monetary reward. American Journal of Mental Deficiency, 1967, 72, 422-427.
Kelleher, R. T. Chaining and conditioned reinforcement. In W. K. Honing (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966, 160-212.
Kratochwill, T. Single subject research. New York: Academic Press, 1978.
Millenson, J. R. Principles of Behavioral Analysis, New York: MacMillan, 1967.
Redd, W. H. The effects of adult presence and stated preferences on the reinforcement control of children's behavior. Merrill-Palmer Quarterly, 1976, 22, 93-98.
Redd, W. H. and Birnbrauer, J. S. Adults as discriminative stimuli for different reinforcement contingencies with retarded children. Journal of Experimental Child Psychology, 1969, 7, 440-447.
Risley, T. R. The effects and side effects of punishing autistic behavior in a deviant child. Journal of Applied Behavior Analysis, 1968, 1, 21-34.
White, O. An investigation of complex chains in the adult, retarded male. Unpublished Master's thesis, University of Oregon, 1970.

The Effect of Reinforcement Schedule on the Intra-Chain Response Patterns of Three Severely Retarded Workers

Robert H. Horner and G. Thomas Bellamy

This research was supported in part by the Bureau of Education for the Handicapped, Research Grant #G00-760-5455.

The Effect of Reinforcement Schedule on the Intra-Chain Response Patterns of Three Severely Retarded Workers

The performance of most vocational tasks requires that a particular chain of responses be performed. Acquisition of the skills to perform a task involve acquisition of each response link within the chain (Bellamy, Horner and Inman, 1979). Following acquisition the entire chain is performed independently. Analysis of task performance by handicapped persons has typically focused on variables affecting the rate or accuracy with which the entire chain is performed (Bellamy, 1975). The experimental research literature, however, argues that even following acquisition there is substantial variability in intra-chain response patterns (Gollub, 1977; Kelleher, 1966; Millenson, 1967). A consistent finding within this literature is that stimulus control at response links is stronger towards the end of an operant chain. As the responder approaches the end of the chain, and the terminal reinforcer which is delivered at chain completion there is typically an increase in response rate (D'Andrea, 1969; Gollub, 1977; Jwaideh, 1973).

The present study provides an application of this experimental research to the field of vocational habilitation of severely retarded adults. The study addresses two questions: (a) what effect does feedback about proximity of reinforcement (i.e., approaching the end of a chain) have on the production rate and intra-chain response patterns of severely retarded workers?; and (b) what effect does the size of the operant chain (i.e. number of response links) have on worker production rate and intra-chain responding?

Method

Subjects. One man and two women employed at the Specialized Training Program, a sub-contract workshop employing severely retarded adults, served as subjects. Subject A, a 37 year old man was functioning at AAMD Adaptive Behavior Level IV at the time of the study. He had been repeatedly labeled profoundly retarded on the basis of standardized IQ tests. He was completely non-verbal and could follow only the simplest of verbal directions.

Subject B was a 21 year old Downs Syndrome woman with a MA of 2.6 years on the Stanford Binet and a Vineland Social Quotient of 2.5 years. Institution records indicated she functioned at AAMD Adaptive Behavior Level IV. She had no expressive language, but was able to follow simple verbal directions.

Subject C, a 31 year old woman, was labeled severely or profoundly retarded on the basis of several standardized IQ tests. She entered the Specialized Training Program workshop approximately two years before the study after spending nine years in a state institution. She has no functional expressive language, but was able to follow a few simple verbal directions.

Setting. The study took place in the workshop of the Specialized Training Program. Subjects performed their tasks with 11 other severely retarded workers in a 290 square foot room. Two supervisors circulated throughout the room providing reinforcement for work completion. This environment has been described in detail by Bellamy, Inman and Horner (1979).

Tasks. The tasks used in the study closely simulated contract tasks available to sub-contract workshops. The tasks performed by Subjects A and B both involved using a plier-like hand tool to cut and crimp the leads of circuit board components. Subject A cut and crimped each lead twice. Subject B cut and crimped each lead only once. Each lead represented a response link in an operant chain defined by the number of components to be completed before a reinforcer was delivered.

Subject C performed a different wire crimping task which required her to place a 16 cm length of wire into a small desktop wire crimping machine. With the wire properly inserted, the subject depressed a handle which crimped the last 4 cm of the wire in a specific multi-figured shape. Each wire served as a response link in the operant chain defined by the reinforcement schedule.

The cycle times for completion of three tasks were relatively short. The crimping done by Subject A required approximately 7 seconds per unit to complete when performed by a non-handicapped worker. The tasks performed by Subjects B and C each required approximately 5 seconds per unit. Due to these short cycle times several units of work were typically completed before a reinforcer was delivered. This pattern is not uncommon in assembly workshops employing both normal and handicapped workers.

Apparatus. A Sodeco printer was used to monitor the time required by each subject to complete a unit of work. The printer was activated when a subject began working and recorded, in tenths of a second, the time to complete each unit. Subject's pressed a button upon completion of each component or wire. When the button was depressed the printer recorded the time between that button pressing response and the previous button pressing response. For Subjects A and B the buttons were placed on their work tables and they were trained to press it just before they set down each finished component. The button for Subject C was attached to the crimping machine, and was automatically depressed each time she lowered the handle.

During those phases of the study where feedback on progress through the chain was unavailable to subjects an automatic counter tracked the number of switch closures. When the chain was completed the counter activated a buzzer. The buzzer served as a discriminative stimulus (S^d) to staff that the schedule was complete, and a reinforcer should be delivered.

Procedure. Each subject was trained on his or her task and performed it for half a day over at least a two week period prior to initiation of the study. During this time, subjects were on the same schedule of reinforcement they experienced during their first phase of the study. When the study commenced, each subject worked on his or her experimental task during one or two of the morning work periods (1 1/2 hours daily). When subjects were not performing experimental tasks they worked on various industry contracts.

The dependent variable throughout the study was the median number of seconds required to complete each unit in the chain. The two independent variables were: (a) manipulation of the number of units a subject needed to complete before delivery of reinforcement (i.e., the Fixed Ratio [FR] schedule defining the chain); and (b) the presence versus absence of extroceptive stimuli providing feedback on how much more work was required before a reinforcer would be delivered. In the presence of extroceptive feedback a subject was on a tandem schedule (Dunham, 1977; Gollub, 1977).

The experimental design of the study assessed the performance of each subject in each of six phases: Chain FR10, Chain FR30, Chain FR60, Tandem FR10, Tandem FR30 and Tandem FR60. The sequence in which subjects received these phases was counterbalanced to provide some control for effects due to order. The counterbalanced order is presented in Table 1. Procedures for each phase were in effect 8-10 days for each subject.

Chained schedules. Under chained schedule phases a subject had access to information on how many units s/he had completed,

TABLE 1

Phases

Subject						
A	Chain FR 10	Chain FR 30	Chain FR 60	Tandem FR 10	Tandem FR 30	Tandem FR 60
B	Tandem FR 30	Tandem FR 60	Tandem FR 10	Chain FR 30	Chain FR 60	Chain FR 10
C	Chain FR 60	Tandem FR 10	Chain FR 30	Tandem FR 60	Chain FR 10	Tandem FR 30

and on how many more were needed before the schedule was complete and a reinforcer would be delivered. This feedback for Subjects A and B was provided via a score card marked with the number of X's required for schedule completion. Each time a unit was completed it was placed on an X. The number of uncovered X's represented the amount of work yet to be done. When all the X's were covered a subject raised his or her hand and a supervisor checked the work and delivered a reinforcer. Subject C received feedback by being given the exact number of wires she had to complete. The closer she came to completing an FR schedule, the fewer wires were lined up in front of her. Each of these feedback systems allowed workers the opportunity to monitor their proximity to reinforcement. Each subject performed his/her task under three chained scheduling, FR10, FR30, and FR60. Eight to ten sessions were conducted under each schedule. The only difference between schedules was the number of units the subjects were required to complete prior to reinforcer delivery.

Tandem schedules. During phases employing Tandem schedules a concerted attempt was made to minimize the information a subject might use as feedback on his/her progress through the chain. For Subjects A and B these involved: (a) making sure the box from which leads were selected contained well over 50,000 leads. In this way the removal of up to several hundred leads had a minimal effect on the perceived quantity in the box. It would, therefore, be very difficult to gain feedback by visually monitoring the number of leads in the box; and (b) once a lead was cut and crimped it was dropped through a small hole into a second box. Completed leads were not visible to the subjects, and hence not available as a source of feedback.

Similar procedures were used with Subject C. During Tandem schedule phases the subject selected wires from a box containing over 300 wires. During all phases she placed crimped wires in an open box also containing over 300 wires. In each case the quantity selected or placed in the boxes was so much less than the total quantity in the boxes that it is unlikely any feedback on progress was available.

Reinforcers. Three types of reinforcers were delivered for work completion: cereals, tokens and staff praise. Dry cereals had been shown to be functional reinforcers for each subject during earlier training efforts. To minimize satiation effects a subject would choose from among several cereals kept in an egg carton at his or her work bench. To balance the total amount of cereal delivered under different FR schedules (i.e. chain length) a subject received one piece of cereal following completion of an FR10 schedule, three pieces following completion of an FR30 schedule and six pieces following completion of an FR60 schedule.

Tokens, in the form of pennies, were delivered in an attempt to maintain consistency with the previous work experience of the subjects. All three subjects were used to receiving a token for work completion. Their low performance levels under dense schedules of token delivery, however, suggests that tokens were minimally functional as reinforcers of work rate. Tokens were exchanged for food at the end of each work period.

Staff contact was delivered only at the completion of an FR schedule. All supervisory staff were instructed not to interact, praise or make eye contact with subjects while they were working. These instructions were designed to ensure that reinforcement of work responses would occur only following the terminal response of the operant chain, and never within the chain.

Results and Discussion

The present study was designed to assess the production rate variability of three severely retarded workers as a function of the type of feedback they received and the fixed ratio schedule used to deliver reinforcement. Of particular interest was a fine grained analysis of intra-chain variability across the six phases. Results will be discussed as they relate to the two major research questions: (a) the effect of feedback about position in the chain (i.e. proximity to a reinforcer), and (b) the effects of fixed ratio schedule size.

The effects of feedback about reinforcer proximity of performance in operant chains. Figures 1 through 9 provide graphic display of intra-chain response patterns for each subject across all six phases (Chain FR10, Chain FR30, Chain FR60, Tandem FR10, Tandem FR30, and Tandem FR60). Each data point is the median number of seconds taken to complete a particular step in the chain. Results for Subject A are provided in Figures 1 through 3. Results for Subject B are provided in Figures 4 through 6, and Subject C's results appear in Figures 7 through 9. Each Figure contains both the tandem and chain performance patterns for one FR schedule.

From Gollub's (1977) recent review of the operant chaining literature it would be expected that regardless of FR size intra-chain responding under tandem schedules would be flat, with all steps being performed in approximately the same amount of time. If subjects did not know when a reinforcer would be delivered no differentiation in responding at specific steps in the chain would be expected. Under a chain schedule, however, subjects' work rates would be expected to improve as they progressed through the chain (i.e., the subjects would be expected to work faster as they approached chain completion and delivery of a reinforcer). In general, these expectations very accurately predicted the response patterns of all three subjects.

With the exception of Subject A's performance under the Tandem FR60 schedule all responding under tandem schedules were remarkably absent of trends. This effect is most aptly represented by the stable Tandem FR30 and Tandem FR60 data of Subjects B and C (see Figures 4, 5, 7, and 8). Although less easily interpreted the same trend stability is evidenced in the Tandem FR10 data for all three subjects (see Figures 3, 6, and 9).

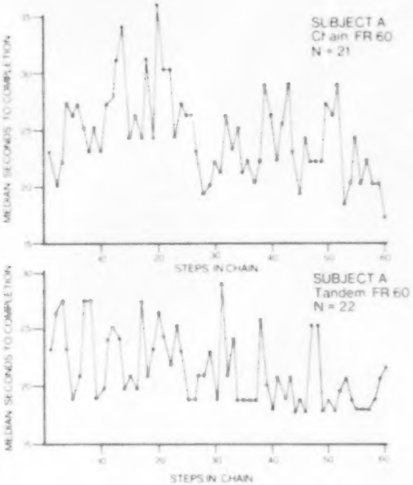


FIGURE 1
Median seconds to completion per step in the chain for Subject A

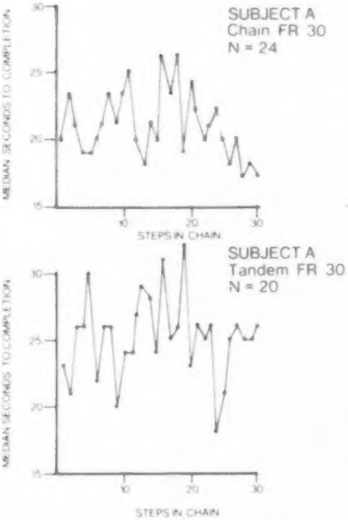


FIGURE 2
Median seconds to completion per step in the chain for Subject A on Chain and Tandem FR30 Schedule

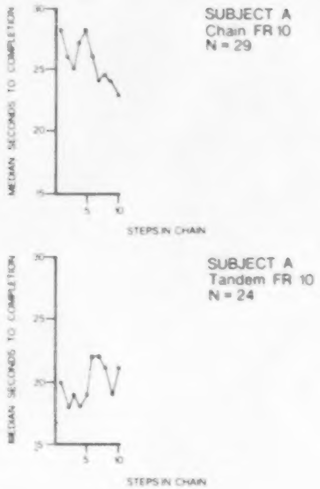


FIGURE 3

Median seconds to completion per step in the chain for Sudject A on Chain and Tandem FR10 Schedules

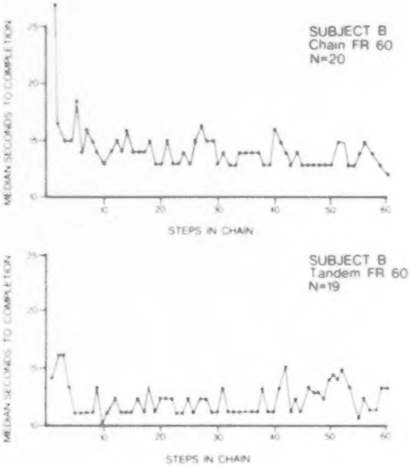


FIGURE 4

Median seconds to completion per step in the chain for Subject B on Chain and Tandem FR60 Schedules

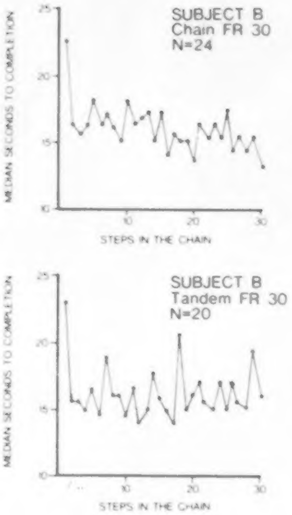


FIGURE 5
Median seconds to completion per step in the chain for Subject B on Chain and Tandem FR30 Schedules

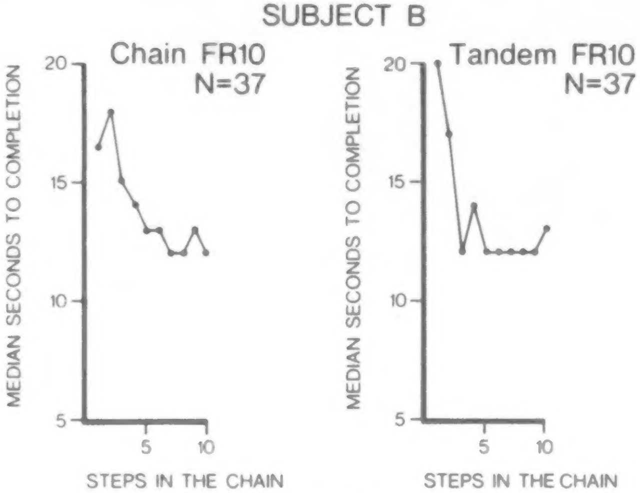


FIGURE 6
Median seconds to completion per step in the chain for Subject B on Chain and Tandem FR10 Schedules

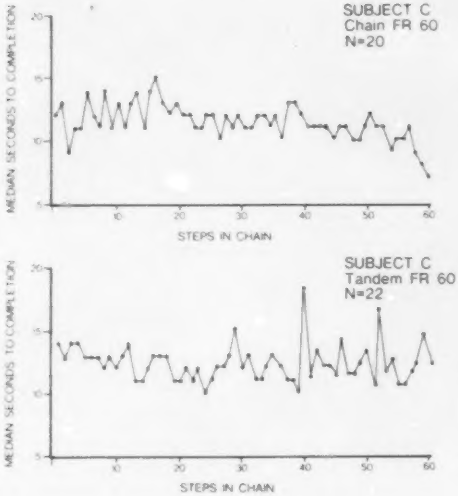


FIGURE 7

Median seconds to completion per step in the chain for Subject C on Chain and Tandem FR60 Schedules

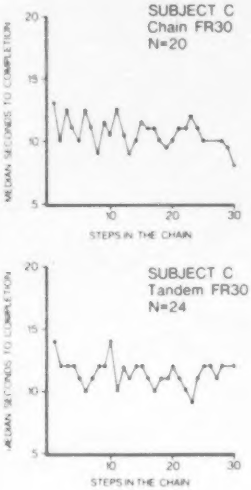


FIGURE 8

Median seconds to completion per step in the chain for Subject C on Chain and Tandem FR30 Schedules

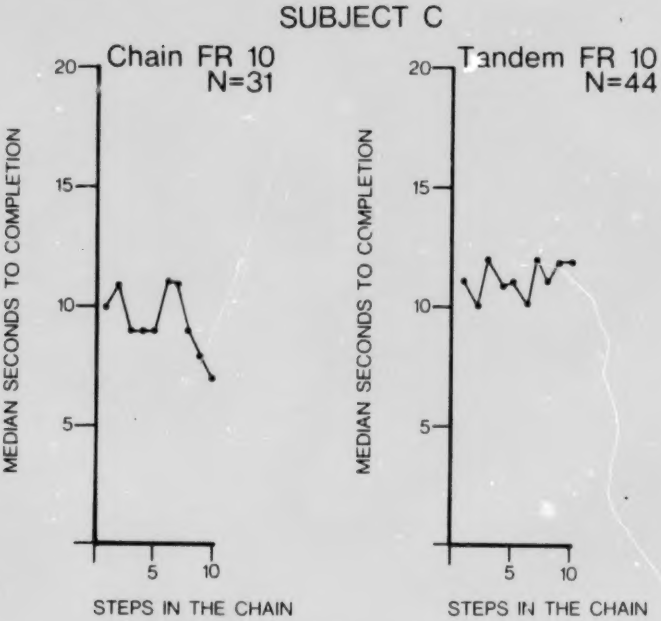


FIGURE 9

Median seconds to completion per step in the chain for Subject C on Chain and Tandem FR10 Schedules

Tandem FR60 and Tandem FR30 results for Subject A (Figures 1 and 2) are less clear with respect to trend stability. Subject A's Tandem FR30 performance indicates substantial variability, and his Tandem FR60 performance suggests a slight improvement in performance as he progressed through the chain. It is of interest to note that the Tandem FR60 phase was the last phase Subject A experienced. It is possible that over the course of the study this subject developed the ability to monitor his progress through the chain without relying on extroceptive feedback.

The performance of all three subjects on Chain FR schedules closely followed the pattern predicted by the experimental literature. Within Figure comparisons of the Chain versus Tandem FR schedules clearly indicate a reduction in response variability and seconds per unit as subjects neared completion of the chain schedules.

As can be seen in Figures 1 through 9 the time required to complete a work unit was consistent only during the first two-thirds of Chain FR schedules. Over the last third there was a predictable and substantial decrease in the variability of responding and the number of seconds taken to complete each of the last few steps of the chain. Without exception, when subjects were on FR schedules they performed the final step of the chain as fast or faster than any other step.

A final observation related to patterns of level and variability of response times focuses on the individual differences that existed between subjects. The performance of Subject C closely fit expected patterns. The behavior of Subject A followed the expected patterns, but with consistently higher levels of variability across all phases. The performance of Subject B, however, followed a consistent post-reinforcement pause pattern (Ferster & Skinner, 1957). After she received a reinforcer, an abnormally long pause would occur before she began the next set of components. This resulted in much longer times for the first step in the chain. Figures 4, 5 and 6 indicated that this pattern is consistent across all FR schedules regardless of whether feedback was delivered (i.e. chain vs. tandem).

Infra-human research of FR schedules has repeatedly documented response patterns in which an abnormally long pause follows reinforcement of an FR cycle. Weiner (1969, 1970) did not find this pattern in his work with normal subjects. The present results suggest that for one subject (Subject C) a pattern of post-reinforcement pausing was predictive of performance across all fixed ratio sizes. This analysis must be made with some reservations, however, given that all schedules in the present study involved fixed ratios. As such, there is insufficient experimental control to demonstrate that the subject's poor performance on the initial steps of a chain were in fact related to fixed ratio schedules and not some other uncontrolled variable.

If variability of intra-chain response patterns is in fact relevant to the vocational habilitation of severely retarded people it should be possible to relate these patterns to overall production rate. In general it would be anticipated that the reduced variability and improved speed on the final steps of chain schedules should make them slightly more efficient than tandem schedules. Results of the mean seconds per unit for each phase offer, at best, equivocal support of this predicted pattern.

The mean seconds to complete a work unit for each subject over each phase are presented in Figure 10. As can be seen from these data Subjects A and C worked slightly faster under Chain FR schedules than under Tandem FR schedules. Subject A averaged 28 sec per work unit across the three chained schedules and 30.5 sec per unit across the three tandem phases. Subject C averaged 12.5 sec per unit under chain schedules and 14.8 sec per unit under tandem schedules.

While these differences are in the predicted direction they are of minimal size, and are not predictive of Subject B's performance. Subject B took an average of 16.3 sec to complete units under chain schedules and only 14 sec per unit during tandem schedules.

Taken together the data from all three subjects suggest that feedback about reinforcement proximity had a substantial effect on intra-chain response rates. The data are less clear, however, with respect to the effect of feedback on overall production rate. Results provided in Figure 10 indicate Subject C's over-

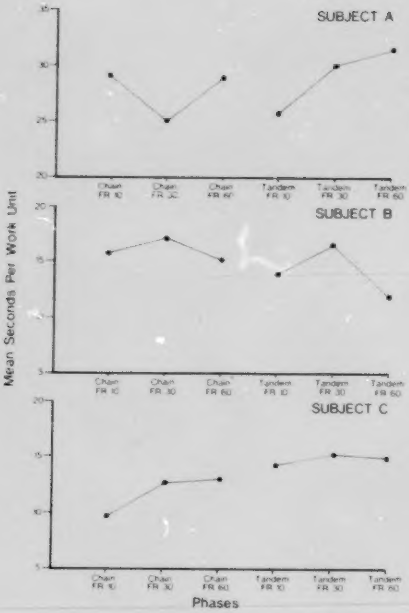


FIGURE 10

Mean seconds per work unit per phase

all production rate was consistently faster when feedback was available and that Subject A worked faster with feedback except on the relatively short FR10 schedule. For Subject B, however, absence of feedback was slightly more effective.

Effects of fixed ratio size. Figure 10 provides the mean number of seconds taken to complete a work unit for each subject across all three FR schedules. Analysis of the previous data provided in Figures 1 through 9 would suggest that chain schedules with small FR sizes would produce faster rates than chain schedules with larger FR sizes, or tandem schedules of any FR size. Chain schedules of small FR size would maximize the frequency with which a subject would rapidly perform the last steps of the chain. This is the opposite effect found by Schroeder (1972), who noted that performance increased as FR size increased.

Results presented in Figure 10 do not indicate a consistent pattern of performance related to FR size. Subject C worked faster during the Chain FR10 phase than any of the other phases. This result fits the expected model suggested by the intra-chain research (Kelleher, 1966) and suggests a possible relationship between Subject C's work rate and fixed ratio size. Results for Subjects A and B, however, provide no replication of Subject C performance patterns. Subject A performed with equal speed during Chain FR10 and Chain FR60 phases, yet worked much faster during the Tandem FR10 phase than during the Tandem FR60 phase. Subject B's best performance occurred during FR60 schedules.

This may, however, have been related to this subject's post-reinforcement pausing pattern. Large FR schedules would decrease the frequency of post-reinforcement pauses. This rationale, however, would also predict that FR10 schedules would be the least effective for this subject. Results provided in Figure 10 do not support this prediction. FR10 schedules were correlated with faster work rates than FR30 schedules in both chain and tandem conditions. Taken together these data do not suggest a functional relationship between the performance rates of Subjects A and B and the size of their fixed ratio schedule of reinforcement.

The present study provides an analysis of the relationship between intra-chain response patterns and the production rate of three severely retarded workers. Results indicate that within a vocational chain the response trend will be reasonably stable if the subject cannot predict when a reinforcer will be delivered. If feedback on reinforcer proximity is provided at each step in the chain, a reduction in response variability will likely occur as the worker proceeds through the task. In addition the final components of the task will be more likely to be performed at an improved rate than if feedback is not provided. This basic pattern exists regardless of FR size.

Implications that these response patterns might have for variability in the production rate of the entire chain are unclear. The variability of subject mean response rates across Chain FR and Tandem FR schedules did not demonstrate a consistent relationship with either FR size or feedback. Many factors could have contributed to these results. It is possible that the unique reinforcement histories of the three subjects may have affected their responsiveness to different FR schedules. Ferster and Skinner (1957) found that the history of an organism could substantially affect that organism's response pattern to different FR schedules. All three of the subjects had worked for over two years in the production setting where the study took place. It is possible that stimuli within the production setting exerted such strong stimulus control over work behavior that the FR and feedback manipulations of the present study were of limited impact.

It is also possible that the reinforcers used in the present study lacked sufficient effectiveness, or did not maintain their effectiveness over time. Spadlin, Girardeau and Corte (1965) found that the performance of profoundly retarded children under FR schedules was affected by the use of low level reinforcers and by "subject saturation". Subject performance patterns under the Chain FR phases suggests that the cereal, praise and tokens were somewhat reinforcing. The lack of an effect between overall production rate and FR size or feedback availability suggests that the effectiveness of these reinforcers may have been minimal.

In general the results emphasize the multiple dimensions that must be considered when evaluating a worker's production rate. Future research will need to look more closely at how these multiple dimensions interrelate. It will be through an understanding of this interrelationship that we will be able to identify the training and supervision procedures most useful for teaching and maintaining the vocational skills of severely retarded adults.

References

Bellamy, G. T. Habilitation of the severely and profoundly retarded: A review of research on work productivity. In G. T. Bellamy (Ed.), Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program, Eugene, Oregon: University of Oregon, Center on Human Development, 1976.

Bellamy, G. T., Horner, R. H., & Inman, D. P. Vocational Habilitation of Severely Retarded Adults: A direct service technology. Baltimore: University Park Press, 1979.

Bellamy, G. T., Inman, D. P., & Horner, R. H. Design of vocational habilitation services for the severely retarded: The Specialized Training Program Model. In G. Hamerlynck (Ed.), Applied behavior analysis techniques for the developmentally disabled. New York: Brunner Mazel, 1979.

D'Andrea, T. Extinction of a heterogeneous chain after several reinforcement schedules. Journal of the Experimental Analysis of Behavior, 1969, 12(1), 127-135.

Dunham, P. The nature of reinforcing stimuli. In W. K. Honig and J. E. R. Staddon (Eds.). Handbook of operant behavior. New Jersey: Prentice-Hall, 1977, p. 98-124.

Ferster, C. B., & Skinner, B. F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.

Gollub, L. Conditioned Reinforcement: Schedule effects. In W. K. Honig and J. E. R. Staddon (Eds.). Handbook of Operant Behavior. New Jersey: Prentice-Hall, 1977, p. 288-309.

Jwaideh, A. R. Responding under chained and tandem fixed-ratio schedules. Journal of the Experimental Analysis of Behavior, 1973, 19, 259-267.

Kelleher, R. T. Chaining and conditioned reinforcement. In W. K. Honig (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966, 160-212.

Millenson, J. R. Principles of Behavior Analysis, New York: MacMillan, 1967.

Schroeder, S. R. Parametric effects of reinforcement frequency, amount of reinforcement, and required response force on sheltered workshop behavior. Journal of Applied Behavior Analysis, 1972, 5(4), 431-441.

Spradlin, J., Girardeau, F., & Corte, E. Fixed ratio and fixed interval behavior of severely and profoundly retarded subjects. Journal of Experimental Child Psychology, 1965, 2, 340-353.

Weiner, H. Controlling human fixed-interval performance. Journal of the Experimental Analysis of Behavior, 1969, 12, 349-373.

Weiner, H. Instructional control of human operant responding during extinction following fixed ratio conditioning. Journal of the Experimental Analysis of Behavior, 1970, 12, 349-373.

Increasing the Productivity of a Profoundly Retarded Sheltered Workshop Employee

Bonnie Biel-Wuerch

This research was supported in part by a contract from the DHEW Bureau of Education for the Handicapped, Program Development Branch.

Increasing the Productivity of a Profoundly Retarded Sheltered Workshop Employee

Reviews on worker productivity (Bellamy, 1976; Bellamy, Inman, & Schwarz, 1978; Gold, 1973; Martin & Pallotta, 1979) suggest there are a variety of procedures effective in increasing and maintaining work rate with retarded subjects. These procedures have been articulated as guidelines for production supervision by several authors (Bellamy, Horner, & Inman, 1979; Martin & Pallotta, 1979) and are being used by many direct service personnel in providing vocational habilitation to severely retarded individuals.

The present study is an addition to this currently available vocational habilitation technology. The focus is on a profoundly retarded man who had demonstrated production of a work task at much higher rates than he was actually performing. Conventional procedures for increasing and maintaining high production rates cited by Bellamy, Horner, and Inman (1979) and Martin and Pallota (1979), including contingent staff contact; social, token, and primary reinforcement for task completion, and verbal prompts to work faster had proved ineffective. This study investigated the utility of an intervention strategy focusing on the details of supervision procedures as a method of determining the critical variables in the maintenance of high levels of productivity.

Method

Subject

A 28 year old male employed at the Specialized Training Program served as the subject in the present study. His selection was based on a history of poor performance rate and failure of traditional production supervision procedures to maintain appropriate work behavior and productivity. At the time of the study he had been employed at the Specialized Training Program (STP) for four years and had been trained to perform several vocational tasks. Prior to his placement in this program the subject had been institutionalized for 22 years. Institution records labeled the subject as profoundly retarded based on an IQ of 10 as measured by the Stanford Binet.

Setting

The study was conducted in the Specialized Training Program at the University of Oregon. The Specialized Training Program is a sub-contract, small parts assembly workshop serving 14 severely and profoundly retarded adults. Workers receive intensive one-to-one training on vocational tasks utilizing procedures described by Bellamy, Horner, and Inman (1979). When a worker demonstrates independent performance on a task to a criterion of 100% accuracy and at a pre-determined rate s/he enters the production room.

The production room is approximately 800 sq. ft. with observational facilities located above and off to one side of the room. Workers sit at individual work stations and independently perform sub-contract work tasks. The production environment is a highly contingent setting. Production supervisors circulate throughout the room providing technical assistance to workers, performing quality control checks, issuing parts, and delivering contingent reinforcement for appropriate work behaviors and task completion. Workers also receive a partial immediate payment in the form of pennies for work completion, which are saved and later used to purchase a variety of edibles during breaks. The work day is comprised of three 1 1/2 hour work sessions. Each work session is followed by a break period; workers leave the production room and have the opportunity to spend money earned during the previous production session.

Data Collection

The dependent variable for the present study was rate of production. This was indexed as the average units per minute that the subject produced during each work session. Data was collected by workshop supervisors who recorded the time the subject began working, the time he stopped working and the number of units produced. Rate for a session was calculated by dividing the total number of units produced by the total number of minutes worked.

Reliability

Reliability was obtained using trained observers who were positioned behind a one-way mirror and collected the same data as production supervisors. All observers were trained to a criterion of 95% observer agreement over two consecutive days prior to the onset of the study.

In determining the reliability of data used to compute rate the crucial factors were (a) the time the subject began work on the task; (b) the time he stopped working on the task; and (c) the number of units produced. Each of these variables was scored for agreement/disagreement by comparing the data recorded by the production supervisors with the data recorded by the reliability observer. Agreement for start-stop times was scored if the supervisor's data was within plus or minus 1 minute of the reliability data. Agreement for units produced was scored if the total number of units produced recorded by the workshop supervisors equalled the total number of units recorded by the reliability observer. Reliability data was collected a minimum of once per phase throughout the study.

The agreement between workshop supervisors and reliability observers was calculated for both start-stop time and units completed. Agreement ranged from 90% to 100% across the study, with agreement for any one phase never averaging less than 91%.

Apparatus

The apparatus used in this study was designed to provide automatic access to music playing on a radio for a duration of 1 minute. Music was selected for use due to its demonstrated reinforcing effect with the subject in a prior study (Horner, 1977). The apparatus was housed in a small wooden box with a coin slot on one side. A small AM/FM transistor radio was wired to the apparatus. To obtain access to music on the radio the subject inserted a penny into the coin slot which in turn activated a timer that turned the radio on for a duration of 1 minute. When the minute had elapsed another penny was required to reactivate the apparatus. The cost of constructing the apparatus was under $50.00.

Task

The subject performed a 5-step vocational task during the study. The task was designed to simulate a wage earning task periodically available to the workshop. The task required the subject to pick up a small wire part, insert it into a fixture, squeeze the handle of the fixture to cut the lead of the part, remove the top portion of the part, and place it in the completed units area. Prior to this study the subject was trained, both to operate the apparatus and perform the work task independently.

Procedure

A single subject ABCBC reversal design (Hersen & Barlow, 1976) was used to evaluate any functional relationship between contingency manipulations and changes in the subject's production rate. Data was collected on worker productivity each session. The study was composed of 5 phases: Baseline, Radio in Break, Radio in Production, Radio in Break, Radio in Production.

Baseline (A). During baseline conditions the subject worked on the work task in the production setting, and was consequated with a penny for completing 4 units. Pennies were saved during a work period and then used to purchase edibles in the break setting. Throughout the study pennies were delivered only for work completion.

Radio in Break (B1). This phase exactly replicated baseline with the exception that pennies earned during a work session could now be used at breaks to purchase 1 minute of music on the radio. The purpose of this phase was to evaluate the effect of changing the backup reinforcer on production rate.

Radio in Production (C 1). During this phase the radio was placed in the production room behind the subject. When the subject earned a penny he was permitted to leave his work station, sit on a nearby chair, and use his penny to activate the radio for 1 minute. When the timer had shut the radio off the subject would return to his work station. The purpose of this phase was to evaluate the effect of an immediate exchange of pennies on production rate.

Radio in Break (B 2). This phase procedurally replicated Radio in Break (B_1). The purpose of this phase was to demonstrate experimental control.

Radio Intervention (C 2). This phase procedurally replicated Radio Intervention (C_1). The purpose of this phase was to reestablish the treatment effect and to provide replication of the experimental effect.

Results and Discussion

Figure 1 presents the mean number of units produced per minute across all phases of the study. Data from phases A, B_1, C_1, B_2, and C_2 indicate increases in work rate and decreases in variability during intervention procedures. Phases B_1, C_1, B_2, and C_2 provide an ABAB reversal design (Hersen & Barlow, 1976) demonstrating this functional relationship.

During phase A, baseline conditions, the subject's performance was highly variable, and work rate ranged from 0 to .27 units per minute. This level is substantially below normal productivity rates. Data indicate that production contingencies in effect during phase A were ineffective in maintaining acceptable performance.

In phase B_1, Radio in Break, the type of reinforcer available to the subject was manipulated. The subject could use pennies earned during production to purchase a high preference reinforcer during break periods. Data indicate an immediate increase in work rate to a level higher than evidenced during baseline conditions. After two sessions of improved performance, work rate gradually decreased across the remaining sessions of the phase, stabilizing at .05 units per minute. Performance during the first two sessions of phase B_1 suggest that the type of reinforcer available may be a relevant contingency variable. However, for this subject, manipulation of this type of reinforcer was not sufficient to result in durable performance.

During phase C_1, Radio in Production, the latency between work completion and reinforcement was the contingency variable manipulated. The subject completed the work task, earned one penny, and immediately exchanged the penny for the high preference reinforcer. Following implementation of C_1 procedures, there was an immediate increase in the level of work rate to .32 units per minute. The average level of performance across this phase was .21 units per minute, higher than performance during baseline and phase B_1. Variability also decreased during phase C_1.

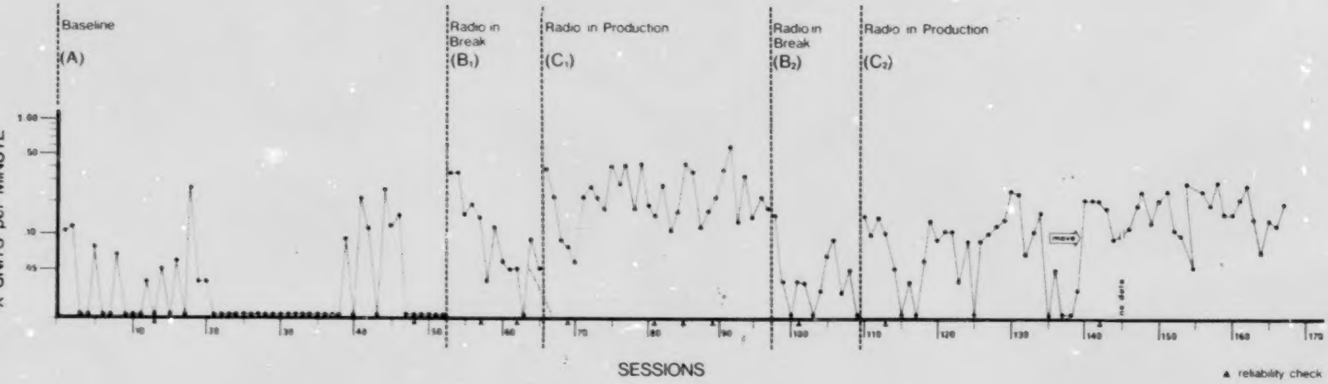


FIGURE 1

The Mean Number of Units Produced per Minute across all Phases of the Study

These data suggest that the latency between work completion and reinforcement is one contingency variable affecting rate. To provide experimental support for this hypothesis a reversal to conditions in effect during phase B_1 was conducted.

Phases B_1, C_1. and B_2 provide an ABA experimental design. During phase B_2 there was an immediate decrease in the level of work rate to a level just below that observed during phase B_1. Data from phases B_1, C_1, and B_2 demonstrate control of increases in rate by the experimental manipulations.

Phase C_2, Radio in Production, represented a return to experimental conditions in effect during phase C_1. Phases B_1, C_1, B_2, and C_2 provide an ABAB reversal design for evaluating the replicability of the effect obtained across phases B_1, C_1, and B_2. Data from phase C_2 indicate an initial increase in work rate to a level equal to that evidenced in phase C_1; followed by 26 sessions of highly variable performance; followed by a more stable performance level approximately equal to that obtained during C_1. Although the average rate of production during phase C_2 closely approximated the level obtained during phase C_1, providing some evidence of an ABAB effect, these data must be interpreted cautiously due to the erratic performance pattern during sessions 110-139.

Reports from production supervisors early in phase C_2 indicate that the implementation or C_2 procedures was closely tied to the addition of a new stimuli to the production environment. Early in phase C_2 a co-worker began directing verbalizations towards the subject. The subject's responses to the co-worker were both extremely disruptive to the workshop and incompatible with task completion. Because of this, the subject's work station was moved increasing the distance between the two workers, and limiting their visual access to one another. (Session 140). Immediately following the move production supervisors reported that the frequency of verbalizations by both workers decreased substantially. Examination of data obtained in C_2 following the move indicate that work rate immediately increased to a level equal to that maintained during phase C_1. In addition, a substantial decrease in variability was noted. This pattern of performance maintained throughout the remainder of the study.

This suggests that the initial portion of phase C_2 did not provide procedural replication of phase C_1 due to the addition of a new stimuli to the production setting (the co-worker's verbalizations). It is quite possible that visual and verbal stimuli provided by the co-worker were more powerful discriminative stimuli for off-task behavior than the work setting was for on-task behavior. When verbalizations were eliminated, with the change in the location of the subject's work station, conditions in effect during the remainder of the phase closely replicated those of phase C_1, and the experimental effect was regained. Of course, lack of experimental design preclude a complete analysis of the functional effect of the addition of the co-worker's behavior to the C_2 conditions.

No data are reported for session #145. The subject was found to be performing the task inappropriately and thus, rate of work completed is not comparable to rates obtained during sessions when the task was performed correctly.

Conclusion

For many retarded workers conventional production supervision procedures are ineffective in maintaining productivity and individualized contingencies must be identified. Data presented illustrate the utility of an intervention strategy focusing on the details of supervision procedures as a method of determining critical contingency variables which influence performance.

The present study demonstrated that the latency between task completion and reinforcement is one contingency variable affecting the productivity of a profoundly retarded subject. These data must be interpreted with caution due to the performance pattern evidenced in phase C_2. In addition, the magnitude of the treatment effect, while a substantial improvement over baseline conditions, is still far below acceptable productivity levels.

Data from phase B_1 suggest that the type of reinforcer selected for use may also be a critical variable. The use of reinforcement to maintain productivity is most likely influenced by multiple parameters of reinforcement including magnitude, duration, and schedule effects. By systematically manipulating these parameters production supervision procedures can be individualized to facilitate maintenance of appropriate work performance.

References

Bellamy, T. Habilitation of the severely and profoundly retarded: A review of research on work productivity. In T. Bellamy (Ed.), Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program. Eugene, Oregon: University of Oregon Center on Human Development, 1976.

Bellamy, T., Horner, R. H., & Inman, D. Vocational habilitation of severely retarded adults. Baltimore: University Park Press, 1979.

Bellamy, T., Inman, D., and Schwarz, R. Vocational training and production supervision: A review of habilitation techniques for the severely and profoundly retarded. In N. Haring and D. Bricker (Eds.), Teaching the severely and profoundly handicapped, Volume 3. Columbus, Ohio: Special Press, 1978.

Gold, M. Research on the vocational habilitation of the retarded: The present, the future. In N. Ellis (Ed.), International review of research in mental retardation (Vol. 6). New York: Academic Press, 1973.

Hersen, M. , and Barlow, D. Single case experimental designs. New York: Pergammon Press, 1976.

Horner, R. H. Stimulus control, transfer and maintenance of upright walking posture with a severely retarded adult. In T. Bellamy, R. Horner and D. Inman (Eds.) Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program. Eugene, Oregon: University of Oregon, 1977.

Martin, G., and Pallotta, A. Behavior modification in sheltered workshops and community group homes for the retarded: Current status and future considerations. In G. Hamerlynck (Ed.) Applied behavior analysis techniques for the developmentally disabled. New York: Brunner Mazel, 1979.

The Effect of Contingent Praise on the Work Rate of a Severely Retarded Adult

Newt Rumble, Dean P. Inman, and Robert H. Horner

University of Oregon

This research was supported in part by a contract from the DHEW Bureau of Education for the Handicapped, Program Development Branch.

The Effect of Contingent Praise on the Work Rate of a Severely Retarded Adult

Within the last two decades there have been repeated demonstrations that severely and profoundly retarded adults are capable of acquiring viable work skills (Bellamy, 1976; Bellamy, Horner & Inman, 1979: Karan, Wehman, Renzaglia, & Schutz, 1976; Martin & Pallotta, 1979). These vocational skills often involved performing remunerative work which, previously, had been available only to "normal" workers who were competitively employed. Although the literature is replete with evidence that the severely retarded can learn work skills, there has not been a massive effort on the part of private enterprise to make remunerative contract tasks available to sheltered workshop programs. Often the work which these programs have been able to obtain represents minimal support for the human effort being undertaken within sheltered workshops. Within the last 20 years researchers and habilitation specialists have been able to secure work contracts for which severely retarded workers were paid to assemble and glue cake boxes (Loos & Tizard, 1955), cut insulated wires to specified lengths (Clarke & Hermelin, 1955), and a myriad of other simple tasks such as folding and stuffing envelopes, collating, label pasting, etc.

Logically, the next step for private industry was to make available more complex and hence, better paying tasks to sheltered workshop programs which serve the severely retarded. A review of the last four years of vocational habilitation literature shows that, while the trend is not yet established nationwide, there is a growing tendency in specified geographic areas for private industry to make highly complex work tasks available to workshop programs. For example, Bellamy and his associate secured remunerative contract work for severely and profoundly retarded adults such as assembly of a 52 piece cam switch activator (Bellamy, Peterson, & Close, 1975), assembly of an 11 wire cable harness which is laced together with 75 self tightening knots (Bellamy, Inman, & Yeates, 1978) placing insulation tape in a protective heat shield for cathode ray tubes (Horner, Lahren, Schwartz, O'Neill, & Hunter, 1979), hand preparation of electronic components and assembly of chain saw nose sprockets and a variety of circuit boards. These impressive vocational gains suggest that it is skill deficits and not retardation per se which have prevented severely retarded individuals from being treated and respected as competent employees.

In view of the above, one might expect that job placements for severely and profoundly retarded adults would have increased greatly in recent years, both within competitive and sheltered settings. This, however, has not been the case. The successes of recent research demonstrating the capacity for severely retarded adults to learn viable vocational skills have not meant that these same adults have had the opportunity to participate in competitive work settings on a long-term basis.

Vocational habilitation specialists are increasingly concerned with finding methods with which to increase the likelihood that severely retarded workers will survive in competitive work settings. Bellamy, Inman, and Horner (1979) have suggested that the employability of a severely or profoundly retarded adult is directly related to two factors. First, the extent to which a retarded worker can perform work to quality control specifications of the customer and second, the extent to which a worker maintains an adequate work rate with minimal supervision. This second factor is the focus of the present paper. Specifically, the paper examines the effects of supervisors delaying their praise on the production rate of one severely retarded adult.

Method

Subject

The subject was a 27-year old woman. She had been labeled severely retarded based on scores obtained from administration of the following assessment tests: Revised Standard Binet, Form L-M (IQ 21); Vineland Social Maturity Scale (mental age 7.2; social quotient, 30); and American Association on Mental Deficiency Adaptive Behavior Scales (scores place subject between Level III and IV). During the study she lived with her natural parents, and was semi-independent (dressed, fed, and groomed herself). She was trained to use community mass transit for travel to and from work, and exhibited limited verbal expression and comprehension.

Setting

The setting for the study was the Specialized Training Program, a subcontract assembly workshop that provides vocational training and production opportunities for severely and profoundly retarded adults. The production setting consisted of 14 severely retarded adult workers (including the subject) who worked at individual work stations on a variety of contract and simulated bench work assembly. The work day consisted of three 1 1/2 hour work periods. During each work period two supervisors circulated around the production floor providing praise for appropriate work behavior, quality controlling completed work, administering various production behavior contingency programs, and providing partial payment for completed work. An in-depth description of this setting is provided by Bellamy, Inman, and Horner (1979).

Task

The subject worked on a task which involved soldering two wire leads to a battery pack, capping each end of the pack with insulated plastic and taping the two caps in place. Production data gathered prior to the study suggested that the subject could work at 60% of the standard rate of non-handicapped workers in industrial settings, but rarely worked at this rate for more than a few moments at a time. At the time of the study the subject had regressed to an average daily work rate that was approximately 30% of norm standard.

Measurement

Task performance was measured on three parameters: (a) minutes taken per unit completed, (b) average daily percent of time on-task, and (c) duration of on-task responses.

Minutes per unit. Productivity data were collected by STP workshop supervisors during each work day. Supervisors recorded the time the subject began working on ten battery packs, and the time when she raised her hand signaling completion of all ten battery packs. From this procedure the number of minutes taken to complete each group of ten battery packs was computed. The average daily minutes per unit score was calculated by dividing the total number of minutes the subject worked in a day by the total number of battery packs she completed.

Percent of time on-task. Data reflecting percent of time on-task also were collected by STP workshop supervisors during all periods of each work day (Bellamy, Horner, & Inman, 1979). On-task behavior was monitored every five minutes throughout each work period. At 5-minute intervals an auditory signal cued a supervisor to record whether the subject was on-task or off-task. At the signal a supervisor observed the subject and scored her as being on-task or off-task. The subject was scored as "on-task" if she was (a) orienting to the task and (b) manipulating task parts to perform one of the steps necessary for task completion.

Duration of on-task responses. The first author observed and recorded the duration (in seconds) of each on-task response during a 10-minute period on seven days in each phase. All duration measures were taken at 11:30 a.m. from behind a one-way mirror which was within 4 feet of the subject's work station. Each time the subject engaged in on-task behavior, the observer started a stopwatch. The watch ran until the subject stopped orienting to the task and/or stopped manipulating parts in a task-appropriate manner. Each duration measure represented the time in seconds during which the subject was on-task.

Reliability

Reliability was calculated for two of the three dependent variables using the percent agreement index. Percent agreement was calculated by counting the number of agreements between observers and supervisors on a recorded score, dividing by the number of agreements plus disagreements and multiplying the outcome by 100.

Minutes per unit. No observer agreement data were taken for this dependent variable.

Percent of time on-task. Observer agreement data for this variable were taken twice with each supervisor per phase. Observer agreement for on-task behavior was recorded on the production floor by an observer and a supervisor. When the five minute timer sounded during the work day, both the observer and supervisor would record the subject as on-task or off-task. Agreement was scored if the observer and supervisor both scored the subject as on-task or both scored the subject as off-task.

Duration of on-task behavior. Observer agreement for on-task behavior was taken simultaneously by an observer and a supervisor during two of the seven 10-minute sessions within each phase of the study (total of eight times). During each session the observer and supervisor sat behind a one-way mirror near the subject and used stopwatches to record the duration the subject stayed on-task each time she initiated an on-task response. Agreement for duration data was defined as time differences within a ±1 second range.

Procedure

A single subject ABAB reversal design was used to investigate the relationship between contingently praising extended durations of on-task behavior and production performance. The four phases of the study were: Baseline I, I-DRE$_5$, Baseline II, and II-DRE$_5$.

Baseline I. During baseline conditions social praise was delivered any time a supervisor observed the subject on-task for longer than a second. Social praise consisted of brief (1-5 second) verbal and/or physical contacts (i.e., "good work") delivered to the subject immediately following the observation of on-task behavior. Physical contacts consisted of light pats, or touching the subject's back, arm, or shoulder and were given only in conjunction with verbal praise. Observations made prior to Baseline I indicated that as supervisors rotated throughout the production room the subject would typically receive social contacts on a schedule of approximately one contact for every 4.8 minutes of the work period. This overall rate of contacts was kept constant across all phases of the study.

Differential reinforcement of extended durations of on-task behavior (I-DRE 5). During the second phase delivery of social praise to the subject was contingent upon supervisors observing 5 consecutive seconds of on-task behavior. Supervisors were instructed to count to five when they observed the subject on-task. If she remained on-task during the entire count they would then approach her and deliver praise. Data were collected for 16 days.

Baseline II. This condition was a replication of the previous Baseline I condition. Data were collected for 32 days.

Differential reinforcement of extended amounts of on-task behavior (II-DRE 5). The conditions which were in effect during the DRE I phase were replicated. Data were collected for 20 days.

Results

Observer agreement was consistently high throughout the study. Agreement for percent time on-task ranged from 71-100% with an average across the study of 95.4%. No phase had an average observer agreement score below 90%. Similarly, observer agreement for duration of on-task responses ranged from 80-100%, with the average for each phase exceeding 90%

To control for the possibility that change in subject behavior was affected by the frequency of supervisor contacts, a counter was kept at the subject's work station and supervisors depressed the counter each time they spoke to, or touched the subject. Data obtained from the counters indicated consistency in supervisor contacts per day across the study. Average daily contact levels across the four phases were 16.06, 18.58, 17.72 and 16.49 respectively.

Minutes per Unit

The average number of minutes per unit per day for each phase is shown in Figure 1. In general the subject took longer

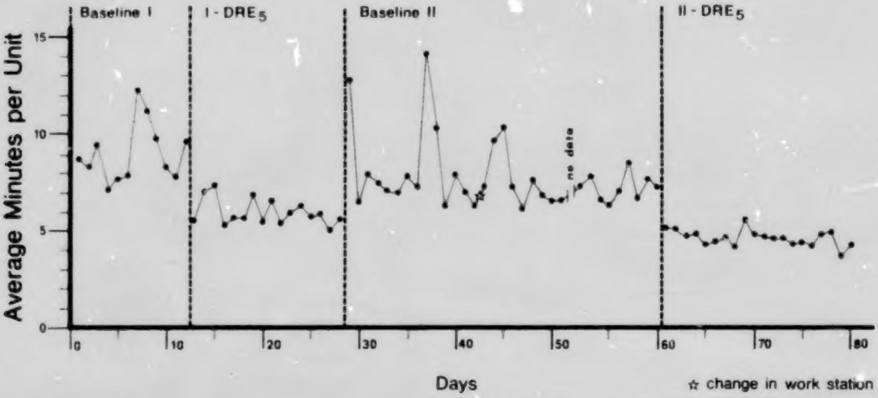


FIGURE 1

Average minutes per unit across work days to complete a unit during Baseline phases than she did during DRE$_5$ phases. Performance within Baseline I averaged 8.97 minutes per unit with a range of 7-12.5. When I-DRE$_5$ began there was an immediate drop in the average minutes to complete a unit, and a slight improving trend that continued across the phase. The mean minutes per unit across I-DRE$_5$ was 1.01 min per unit (11%) improvement over Baseline I. In addition, performance across I-DRE$_5$ was much more stable than that observed during Baseline I (range 5.0-7.3).

Subject performance during Baseline II was similar to that observed during the Baseline I condition, in both level and variability. An immediate increase in minutes to completion (from 5.64 minutes per unit to 12.85 minutes per unit) was noted on the first day of the Baseline II. The mean time per unit for the phase was 7.8; nearly 2 min per unit slower than I-DRE$_5$ performance, but not a complete recovery to the Baseline I level. Scores during Baseline II ranged from 6.3 to 12.4 minutes per unit.

During the final phase of the study (II-DRE$_5$) the subject performed faster and more consistently than at any previous point. Her mean performance across the phase was 4.4 minutes per unit within a 2 minute range.

Percent of Time On-Task

As can be seen in Figure 2 the subject's percent of time on-task followed a pattern similar to that observed with the

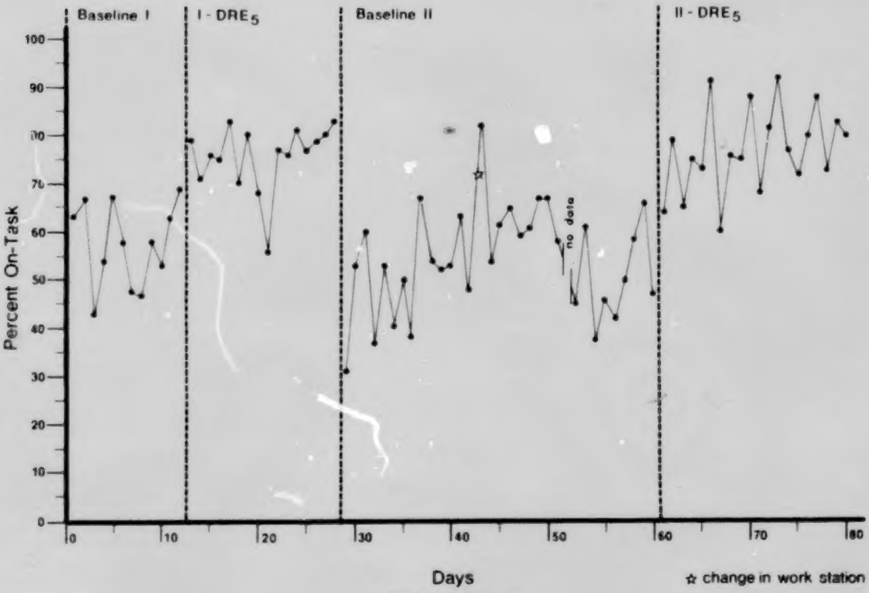


FIGURE 2

Percent time on-task across work days

Minutes per Unit data. Baseline phases were characterized by low levels of on-task behavior with immediate and sustained improvements occurring when DRE_5 procedures were introduced. The mean percent of on-task behavior was 57.59%, 75.69%, 56% and 76.55% respectively for the four phases.

Duration of On-Task Responses

Figure 3 indicates the length or duration of the subject's on-task responses for each phase. The abscissa for each of the four graphs denotes the duration (in seconds) of on-task behavior. The ordinate reflects the percent of total on-task responses across the phase for each duration. As can be seen in Figure 3, Baseline conditions were characterized by a skewed distribution toward responses of 10 seconds or less in duration while DRE_5 conditions evidence a flattening of the distribution, indicating that a greater percent of the subject's on-task responses lasted 10 seconds or more.

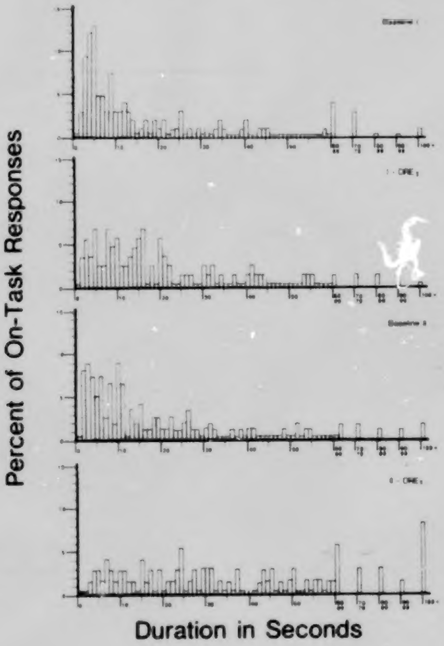


FIGURE 3
Duration of on-task responses in seconds

During the Baseline I condition, 58.6% of all on-task responses were less than or equal to 10 seconds in duration. This indicates a strong tendency toward on-task responses which were so short that they precluded much opportunity for the subject to complete her work.

When the I-DRE$_5$ condition was in effect, a flattening of the distribution is noted. During this phase only 38.7% of the responses were less than or equal to 10 seconds in duration, and a greater percent of longer on-task responses occurred. Of the total responses for the phase, 38.7% were less than or equal to 10 seconds in duration (as compared to 58.6% for Baseline I).

The Baseline II phase was characterized by a skew in the distribution very similar to that observed for the Baseline I phase. During this phase 53.1% of all on-task responses were less than or equal to 10 seconds in duration.

During II-DRE$_5$, the flattened distribution associated with the I-DRE$_5$ phase was again observed. However, as seen in Figure 3, the distribution is flatter at the left end of the scale, representing short durations of on-task responding, with more responses at the opposite end of the scale (greater than or equal to 60 seconds).

Only 18.2% of all responses were less than or equal to 10 seconds in duration during II-DRE$_5$ (compared with 58.6%, Baseline I; 38.7%, I-DRE$_5$; and 53.1%, Baseline II) with 19.5% of responses greater than 60 seconds (compared with 7.3, Baseline I; 4.4, I-DRE$_5$; and 6.1 for Baseline II). In general a comparison of the duration data indicates a classic ABAB effect across the four phases of the study.

Discussion

The results indicate that making supervisor praise contingent upon an increased duration of on-task behavior had a positive effect on subject productivity. The subject produced units in less time, was generally on-task more and when she went on-task remained there longer when DRE procedures were used. The use of an ABAB reversal design demonstrates that these positive changes were functionally related to the intervention procedure.

When DRE procedures were used the subject received praise only when she was on-task for 5 seconds or more. Three characteristics of this intervention are important. First, the procedure systematically excluded the delivery of praise for on-task responses of less than 5 seconds (38% of the subject's Baseline I responses), while providing no upper limit on how long the response might last. Second, the use of a 5 second contingency may have decreased the subject's ability to predict when praise would be delivered. During Baseline phases praise was more consistently correlated with presence of a supervisor. Hence the subject could track supervisor location as a method of anticipating when on-task behavior might be rewarded. The DRE procedure resulted in supervisor presence being less highly correlated with praise.

A final characteristic of the DRE procedure was its ease of implementation. Production supervisors were trained to use the procedure in one 15 min session, and were able to easily switch from Baseline to DRE contingency rules at phase change points. Thus, an easily implemented change in supervision behavior had a major change in the vocational performance of the subject.

The present study demonstrates that extending the duration of on-task behavior required for praise can have a substantial impact on the behavior of severely handicapped workers. This procedure may also be functional in non-work settings. Classrooms, homes, and communities all require individuals to perform on-task behavior. As such the present delay procedure may be a useful addition to the programming strategies used by teachers, group home staff or any service providers working with severely retarded people.

References

Bellamy, G. T. (Ed.). *Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program*. Eugene, Oregon: University of Oregon Center on Human Development, 1976.

Bellamy, G. T., Horner, R. H., & Inman, D. P. (Eds.). *Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program*, Vol. 2. Eugene, Oregon: University of Oregon Center on Human Development, 1979.

Bellamy, G. T., Inman, D. P., & Horner, R. H. Design of vocational habilitation services for the severely retarded: The Specialized Training Program Model. In G. Hamerlynck (Ed.), *Applied behavior analysis techniques for the developmentally disabled*. New York: Brunner Mazel, 1979.

Bellamy, G. T., Inman, D. P., & Schwarz, R. Vocational training and production supervision: A review of habilitation techniques for the severely and profoundly retarded. In N. Haring and D. Bricker (Eds.), *Teaching the severely and profoundly handicapped*, Vol. 3. Columbus, Ohio: Special Press, 1978.

Bellamy, G. T., Inman, D. P., & Yeates, J. Workshop supervision: Evaluation of a procedure for production management with the severely retarded. *Mental Retardation*, 1978, *16*, 317-319.

Bellamy, T., Peterson, L., & Close, D. Habilitation of the severely and profoundly retarded: Illustrations of competence. *Education and Training of the Mentally Retarded*, 1975, *10*, 174-186.

Clarke, A., & Hermelin, F. Adult imbeciles: Their abilities and trainability. *The Lancet*, 1955, *2*, 337-339.

Horner, R. H., & Bellamy, G. T. Long term structured employment: Productivity and productive capacity. In G. T. Bellamy, G. O'Connor, and O. C. Karan (Eds.), *Vocational Rehabilitation of Severely Handicapped Persons: Contemporary service strategies*. Baltimore: University Park Press, 1979.

Horner, R. H., Lahren, B., Schwartz, T., O'Neill, C. T., & Hunter, J. D. Considerations in dealing with low production rates of severely retarded workers. *AAESPH Review*, 1979, *4*, 202-212.

Karan, O., Wehman, P., Renzaglia, A., & Schutz, R. *Habilitation practices with the severely developmentally disabled*. University of Wisconsin, Waisman Center on Mental Retardation and Human Development, Madison, Wisconsin, 1976.

Loos, F. M., & Tizard, J. The employment of adult imbeciles in a hospital workshop. American Journal of Mental Deficiency, 1955, 59, 395-403.

Martin, G., & Pallotta, A. Behavior modification in sheltered workshops and community group homes for the retarded. In G. Hamerlynck (Ed.), Applied behavior analysis techniques for the developmentally disabled. New York: Brunner Mazel, 1979.

Systematic Correction Procedures in Vocational-Skill Training of Severely Retarded Individuals

Daniel W. Close, Larry K. Irvin
Herbert J. Prehm and Valerie E. Taylor

University of Oregon

This report originally appeared in the American Journal of Mental Deficiency, 1979, 83, 270-275. Reprinted with permission.

Systematic Correction Procedures in Vocational-Skill Training of Severely Retarded Individuals

Much research has focused on factors that influence vocational skill acquisition by severely retarded persons because of the apparent reduced ability of these individuals to select and attend to the relevant dimensions in learning tasks (Fisher & Zeaman, 1973; Gold & Scott, 1971). Studies which derive from Zeaman and House's (1963) Attention Theory have demonstrated that stimulus manipulation procedures function to direct severely retarded learners' attention to the relevant dimensions of discriminations which comprise complex vocational tasks (Gold, 1972; Irvin & Bellamy, 1977). Additional investigations which focus on the relative efficacy of systematic verbal and physical correction procedures by trainers are also necessary. Because retarded persons appear to form avoidance tendencies more slowly than approach responses in the initial stages of discrimination learning (Zeaman & House, 1962), it is clear that correction of errors is an important facet of discrimination training with these individuals. In addition, such trainer-related correction methods minimize the need for the types of human, time, and material resources required by stimulus-related procedures (color-coding, or initial presentation of stimuli with exaggerated S+/S- cue differences, and subsequent fading).

Until recently, errors were treated as a problem of extinction in studies of the learning ability of severely and profoundly retarded individuals (Bensberg, Colwell, & Cassel, 1965). This method of training appears to be inadequate for non-retarded as well as retarded persons, however, because little, if any, information about error behavior is provided (Spence, 1964; Zeaman & House, 1962).

Training procedures that employ systematic feedback following errors have been and continue to be developed for use with severely retarded individuals. Such methods involve the use of verbal correction (Bellamy, Peterson, & Close, 1975; Butterfield & McIntyre, 1969; Butterfield & Zigler, 1965; Gold, 1972; Gold & Barclay, 1973; Irvin & Bellamy, 1977), physical correction (Bellamy, et al., 1975; Bensberg, Colwell, & Cassel, 1965; Gold, 1972; Gold & Barclay, 1973; Irvin &

Bellamy, 1977), and repeated practice or "overcorrection" (Azrin, Kaplan, & Foxx, 1973; Azrin & Wezolowski, 1974; Bellamy, et al., 1975; Rusch, Close, Hops, & Agosta, 1976) for discrimination and manipulation errors in skill training with severely retarded persons.

The facilitative effects of these systematic correction procedures in teaching complex skills have been demonstrated by such prior research. Attempts have not been made, however, to assess the relative effectiveness of the various correction procedures, i. e., systematic manipulation of combinations of verbal and physical correction procedures. This issue is of some importance to trainers of severely retarded individuals; human and other resources available to teach functional skills to such persons are typically limited. As a result, severely retarded employees in sheltered workshops are often confined to work on simple, repetitive, poorly paying tasks because of an assumed inability to learn and perform complex tasks. Only with empirical information can data-based decisions be made regarding the methodology of choice for training severely retarded persons on more habilitative, difficult and/or complex vocational skill tasks.

The current study was designed to address two relative efficacy issues regarding vocational skill training of severely retarded persons: the effects of different types of verbal correction and the effects of different types of systematic physical correction on severely retarded individuals' performance of a difficult discrimination task.

Method

Design

The two independent variables were verbal correction and systematic physical correction. Two types of verbal correction (general verbal cue--"Try another way," and specific verbal cue--"Flat side in") and three types of systematic physical correction procedure (gesture, physical prompt, and repeated practice) were crossed to yield six between-subject treatment groups.

Subjects

Seventy severely retarded adults residing in two private residential facilities were randomly selected as participants in the study. Selection criteria were: IQ/SQ in the 20-37 range; chronological age from 16 to 45 years; no visual or auditory disabilities; and, no manual handicap, extreme disruptive or physically abusive behavior that prevented performance of the criterion task. The subjects were randomly assigned to the six treatment conditions; 12 were assigned to five of the groups, and ten to one group. Subject IQ/SQ and age characteristics were comparable for the six groups; mean IQ/SQ was between 27 and 30. Mean age in years was between 26 and 31 for each group. IQ's were determined from either the Stanford-Binet Intelligence Scale or the Wechsler Adult Intelligence Scale. The SQ scores were derived from the Vineland Social Maturity Scale.

Ten subjects were eliminated during the study due to agitated and disruptive behavior that, to decrease, would require longer term treatment than the training sessions allowed. Additional subjects were randomly selected to replace the eliminated subjects.

Material and Apparatus

The materials for the criterion task were 10 cm diameter Phillips bicycle front axle nuts and 7 cm diameter, 50 cm long, round head machine screws comparable to bicycle axle posts. One face of the axle nut was flat and the other raised 1 mm. Ten nuts and bolts were used.

The materials were placed in a small wooden bin with two compartments separated by a space in the middle. The open sides of the tray faced the subject. The compartment to the subject's left contained the 10 Phillips axle nuts. The compartment to the right contained the 10 round head machine screws. The space between the compartments served as the repository for completed tasks.

Experimenters

Two experienced trainers who were graduate students in special education served as experimenters. To assure fidelity of treatment implementation, inter-experimenter reliability was assessed for ten minutes per day during each day training was conducted. There was 100% agreement between measures collected by the experimenters during each reliability check.

Procedure

Subjects were trained individually while seated at a table facing the apparatus. Each training session lasted 50 minutes, 100 trials, or until criterion performance of task completion, whichever occurred first.

Completion of the task required the following sequence of movements; (a) reach into the left compartment with the left hand, grasp one axle nut with thumb and index finger, and withdraw it from the apparatus; (b) look at the axle nut and rotate it with the right thumb and index finger until the raised side faces up; (c) reach with the right hand into the right compartment, grasp one axle post, and insert it into the flat side of the axle nut; and (d) place completed assembly between the compartments.

Familiarization

The first stage of training was used to familiarize each subject with the materials, the movements required by the task, and the verbal corrections to be used by the experimenter during training. During familiarization, the criterion discrimination was not taught or performed. The trainer demonstrated the entire sequence of movements, without highlighting the difficult discrimination in any way, while the subject observed. The trainer then said, "You do it" and pointed to the left container. If the subject performed the required movements of placing the bolt onto the nut for five consecutive trials, without regard for the discrimination of interest, familiarization training was concluded. If the subject failed to initiate the task or performed an incorrect movement, the trainer physically guided the subject's hands through the required movements until five consecutive correct, independent trials were achieved. Verbal feedback such as "Good" was provided by the trainer for each correct movement.

Training

Once the subject correctly performed the movements required by the task, training on the criterion discrimination task was initiated. Before the first trial on the training task, a demonstration procedure was conducted by the experimenter. In the demonstration, the experimenter performed both the correct and incorrect methods of completing the task. Purposeful errors during this demonstration were corrected as follows. Subjects who were to be trained with a verbal correction of "Try another way" were presented with that verbal correction plus a gesture directed at the incorrect position of the nut in the experimenter's hand. The experimenter then positioned the nut correctly in his/her own hand and said, "Good." Subjects who were to be trained using a verbal correction of "Flat side in" were presented with that verbal correction plus a gesture directed at the incorrect position of the nut in the experimenter's hand. The experimenter then positioned the nut correctly in his/her hand and said, "Good." The demonstration procedure was presented five times for all subjects. The purpose of the demonstration was to show the subject how to respond to the verbal corrections to be used, rather than to teach the criterion discrimination.

Following the demonstration, the trainer said, "You do it" and pointed to the left container. The subjects received presumed reinforcers such as verbal compliments and gentle physical contact following correct performance of the discrimination task.

A correction procedure that varied by training group followed any discrimination error or failure to initiate the task. When a correction was used, the trial was scored as incorrect and the subjects' response to the correction was not scored. Correction was provided as follows:

> **'Flat side in' -- Gesture condition.** The trainer said, "Flat side in" and gestured to the incorrect position of the nut. If the subject did not correct the error within five seconds the trainer removed the nut from the subject's hand and placed it in the left bin. The trainer then gestured or prompted the subject to reach into the left bin and attempt the next trial.

'Try another way' -- Gesture condition. The trainer said, "Try another way" and gestured to the incorrect position of the nut. If the subject did not correct the error within five seconds the trainer removed the nut from the subject's hand and placed it in the left bin. The trainer then gestured or prompted the subject to reach into the left bin and attempt the next trial.

'Flat side in' -- Physical prompt condition. The trainer said, "Flat side in" and, using a right hand pincer grasp, positioned the nut correctly in the subject's left hand, saying, "Good." The subject then completed the assembly, with physical assistance if necessary. The trainer gestured or prompted the subject to reach into the left bin if no initiation to the next trial was made within five seconds of the completion of the assembly.

'Try another way' -- Physical prompt condition. The trainer said, "Try another way" and, using a right hand pincer grasp, positioned the nut correctly in the subject's left hand, saying, "Good." The subject then completed the assembly, with physical assistance if necessary. The trainer gestured or prompted the subject to reach into the left bin if no initiation to the next trial was made within five seconds of the completion of the assembly.

'Flat side in' -- Repetition of physical prompt five times. The trainer said, "Flat side in" and physically prompted the correct response, saying, "Good." The trainer then removed the nut from the subject's hand, placed it flat side up in the bin, and assisted the subject in picking up the nut. With the nut in the incorrect position in the subject's hand, the trainer said, "Flat side in" and physically prompted the correct positioning of the nut in the subject's hand. This correction procedure was repeated a total of five times for each initial discrimination error. The nut and bolt were assembled by the subject after the fifth repetition of the physical prompt, with physical assistance from the trainer if necessary. A trial was terminated and a new trial initiated if the subject independently corrected an error during the repetition of the physical prompt correction procedure.

'Try another way' -- Repetition of physical prompt five times. The trainer said, "Try another way" and physically prompted the correct response, saying, "Good." The trainer then removed the nut from the subject's hand, placed it flat side up in the bin, and assisted the subject in picking up the nut. With the nut in the incorrect position in the subject's hand, the trainer said, "Flat side in" and physically prompted the correct positioning of the nut in the subject's hand. This correction procedure was repeated a total of five times for each initial discrimination error.

The nut and bolt were assembled after the fifth repetition of the physical prompt with physical assistance from the trainer if necessary. A trial was terminated and a new trial initiated if the subject independently corrected an error during the repetition of the physical prompt correction procedure.

Twelve out of fourteen correct discriminations and placement of the axle nut on the axle post constituted criterion acquisition.

Results

The independent variables in the study were verbal correction and systematic physical correction. Discrimination learning was measured by the number of trials to the acquisition criterion. The data were analyzed using a 2 x 3 (type of verbal correction x type of systematic physical correction) factorial analysis of variance. Table 1 shows the descriptive statistics for trials to criterion acquisition. Subjects who did not reach criterion acquisition were terminated after 200 trials and given a score of 200.

TABLE 1

Group Means and Standard Deviations for Trials-to-Criterion

		GESTURE	PROMPT	REPEATED PRACTICE
'Try Another Way'	$\overline{X}$	107.75	55.92	23.00
	S.D.	22.72	8.11	2.70
'Flat Side In'	$\overline{X}$	94.50	67.00	32.58
	S.D.	20.84	10.69	4.89

There was a reliable effect due to type of systematic physical correction procedure ($F = 13.92$, 2/64 df, $p < .01$). No significant effect due to verbal correction was apparent, and no significant interaction effect occurred.

Scheffe's multiple means comparison tests revealed that: (a) subjects in the gesture group required more trials than subjects in the physical prompt condition ($p < .01$); and (b) subjects in the gesture and physical prompt conditions required more trials than subjects in the repeated correction condition ($p < .001$, gesture; $p < .05$, physical prompt).

Discussion

The findings help clarify the role of verbal statements in training severely retarded persons on vocational skill tasks involving difficult visual-motor discriminations. Gold (1972) demonstrated that general verbal assistance ("Try another way") was effective for correcting discrimination errors by severely retarded adults on a complex assembly task. The results of the present study can be interpreted as indicating that a verbal statement, regardless of the specificity of its content, functions to inform the severely retarded learner that an error has been made. This finding is consistent with Butterfield and McIntyre's (1969) results with mildly retarded subjects on concept-switching tasks.

More than verbal assistance appears to be needed, however, to teach severely retarded individuals the precise movements to correctly perform vocational skill tasks. Differential pairing of verbal and systematic physical correction procedures appears to produce varying degrees of "attention" by severely retarded learners to the specific behavior(s) to be performed. These results extend previous findings in one additional way by providing a direct comparison between repeated practice and traditional physical prompting procedures. The bulk of the relevant literature details the use of simple physical prompting procedures to teach vocational and self-care skills to severely retarded persons (Bensberg, Colwell, & Cassel, 1965; Gold, 1972; Irvin & Bellamy, 1977; Lent, 1975). The results of the current study illustrate the applicability of repeated practice or "overcorrection" procedures for teaching vocational skills that require difficult visual-motor discriminations. Repeated practice produced substantially and significantly more rapid acquisition of the criterion task than did simple physical prompting procedures.

The relative superiority of the repeated practice procedure for training difficult visual-motor discrimination tasks also has some important implications for theories regarding learning by retarded individuals. Fisher and Zeaman's (1973) attention-retention theory postulates that retarded persons have a deficit in perception of relevant aspects of visual discriminations. They and others (Gold & Scott, 1971) assert that training procedures should be designed to increase the saliency of relevant visual cues for retarded learners. As a result, much of the research that has focused on application of this theoretical framework employs a strategy that involves the manipulation of stimulus dimensions to engineer learning by retarded individuals (Gold, 1972; Gold, 1973; Gold & Barclay, 1973; Irvin & Bellamy, 1977). The findings in the present study, however, demonstrate that trainer behavior can also be manipulated systematically to produce significant gains in criterion acquisition. This result is especially noteworthy because the utility of attention-engineering strategies in which trainer or response cues are manipulated may often be more functional than that of other approaches. Such trainer-related procedures are applicable across a wide variety of tasks, whereas, with many tasks, it may not be practical or feasible to employ stimulus manipulation training procedures.

Conclusions in the present and previous research (Gold, 1972; Irvin & Bellamy, 1977) may not be applicable to a small group of difficult-to-train adults. Nine subjects in the physical prompt and repeated practice groups did not reach acquisition due to overt resistance to training. Further research is clearly needed to determine under which circumstances the effects of these highly efficient training procedures are predictable. The result will be a better understanding of the role of trainer behavior in the acquisition of difficult discriminations by severely retarded persons.

References

Azrin, N. H., Kaplan, S. J., & Foxx, R. M. Autism reversal: Eliminating stereotyped self-stimulation of retarded individuals. *American Journal of Mental Deficiency*, 1973, *78*, 241-248.

Azrin, N. H., & Wezolowski, M. D. Theft reversal: An overcorrection procedure for eliminating stealing by retarded persons. *Journal of Applied Behavior Analysis*, 1974, *7*, 577-581.

Bellamy, T., Peterson, L., & Close, D. Habilitation of the severely and profoundly retarded: Illustrations of competence. *Education and Training of the Mentally Retarded*, 1975, *10*, 174-186.

Bensberg, G. J., Colwell, C. N., & Cassel, R. H. Teaching the profoundly retarded self-help activities by behavior shaping techniques. *American Journal of Mental Deficiency*, 1965, *69*, 674-679.

Butterfield, E., & McIntyre, A. Cognitive and motivational factors in concept switching among the retarded. *American Journal of Mental Deficiency*, 1969, *74*, 235-241.

Butterfield, E. C., & Zigler, E. The effects of success and failure on the discrimination learning of normal and retarded children. *Journal of Abnormal Psychology*, 1965, *70*, 25-31.

Crosson, J. The functional analysis of behavior: A technology for special education practices. *Mental Retardation*, 1969, *7*(4), 15-19.

Fisher, M. A., & Zeaman, D. An attention-retention theory of retardate discrimination learning. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 6). New York: Academic Press, 1973, pp. 169-256.

Foxx, R., & Azrin, N. The elimination of autistic self-stimulatory behavior by overcorrection. *Journal of Applied Behavior Analysis*, 1973, *6*, 1-14.

Gold, M. Research on the vocational habilitation of the retarded: The present, the future. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 6). New York: Academic Press, 1973, pp. 97-147.

Gold, M. Stimulus factors in the skill training of the retarded on a complex assembly task: Acquisition, transfer, and retention. *American Journal of Mental Deficiency*, 1972, *76*, 517-526.

Gold, M., & Barclay, C. The learning of difficult visual discriminations by the moderately and severely retarded. Mental Retardation, 1973, 11(2), 9-11.

Gold, M., & Scott, K. Discrimination learning. In W. B. Stephens (Ed.), Training the developing young. New York: John Day, 1971, pp. 420-444.

Irvin, L. K., & Bellamy, T. Manipulation of stimulus features in vocational skill training of severely retarded individuals. American Journal of Mental Deficiency, 1977, 81, 486-491.

Lent, J. R. Developing daily living skills for the mentally retarded. In J. M. Kaufman & J. S. Payne (Eds.), Mental Retardation: Introduction and personal perspectives. New York: Charles E. Merrill, 1975.

Rusch, F., Close, D., Hops, H., & Agosta, J. Overcorrection: Generalization and maintenance. Journal of Applied Behavior Analysis, 1976, 9, 498.

Spence, J. Verbal discrimination performance under different verbal reinforcement combinations. Journal of Experimental Psychology, 1964, 67, 195-197.

Zeaman, D., & House, B. J. Approach and avoidance in the discrimination learning of retardates. Child Development, 1962, 35, 355-372.

Zeaman, D., & House, B. J. The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), Handbook of mental deficiency. New York: McGraw-Hill, 1963, pp. 159-166.

The Effects of Immediate Correction on the Error Rate of One Severely Retarded Worker

Robert H. Horner and Nancy M. Prill

This research was supported in part by the Bureau of Education for the Handicapped, Research Grant #G00-760-5455.

The Effects of Immediate Correction on the Error Rate of One Severely Retarded Worker

Response rate and response accuracy are important concerns in many instructional environments. This is particularly true in a workshop environment where the speed and errorlessness with which products are constructed is directly related to the economic viability of both individual workers and the entire workshop (Federal Register, 1974). One variable which has been shown to affect both rate and accuracy is the latency between a response and its consequence. Both experimental (Catania, 1970, Jenkins, 1970) and applied (Fink & Carnine, 1975, Guralnick, 1976) research has found that the time delay between a response and its consequence affects the rate and accuracy of a response. The present study extends this research to decisions made by workshop supervisors in a production environment. The specific questions addressed are: (a) does the latency between performance of a response and consequation of the response affect the likelihood of response accuracy, and (b) does the latency between performance of a response and consequation of the response affect the time taken to perform the response.

METHOD

Subject and setting

The subject was a 29-year old Down's Syndrome woman. Her scores on the Stanford Binet and Adaptive Behavior Scale indicated profound retardation. She exhibited no major motor or sensory impairments. The subject was non-verbal, but was able to use a limited number of ASL signs, and could imitate trainer behavior.

At the time of the study, the subject was employed at the Specialized Training Program, a subcontract assembly shop employing fourteen severely retarded adults. Each worker in this environment spends three, 1 1/2 hour work periods performing small parts electronics assembly each day. Two production supervisors rotate around the room providing task assistance, checking the quality of products, praising workers for on-task behavior and reinforcing work completion. This environment has been described in detail by Bellamy, Inman and Horner (1979).

Task

The task involved peeling a waxed paper backing from each of four .63 cm x 3.3 cm pieces of cushioning tape and placing one tape in each corner of a rectangular-shaped metal heat shield. Quality control specifications required that the tape be centered exactly in the corner of the shield, and be no more than .79 cm nor less than .32 cm from the top edge of the shield.

Observation

The dependent variables for the study were the number of minutes taken to complete a shield and the percent of tapes placed incorrectly. Observation of the minutes to complete a shield was performed by workshop supervisors who wrote down the time the subject began working on a shield and the time when she completed the shield. From this data a daily summary score was calculated by taking the total number of minutes worked on shields divided by the total number of shields produced. Time the subject spent correcting errors was not included in the total minutes worked on shields.

Errors were recorded only for the initial placement of each tape in the shield. Tape placement was assessed by supervisors using a measuring stick to determine the accuracy of tape depth, and their own judgement to assess whether the tape was in the shield corner. Any error resulted in the tape being scored as incorrectly placed. The type of error was not scored. Error data was summarized daily by taking the total number of tapes placed incorrectly over the total number of tapes placed. This figure provided a percent of errors in tape placement per day.

Reliability

Reliability for error data was measured on the production floor by having two supervisors independently evaluate tape placement. A percent agreement figure was then calculated by taking the total number of tapes in which both supervisors agreed that the tape was correct or incorrect over the total number of tapes evaluated by two supervisors. These data were collected only after all four tapes had been placed in a shield. During intervention phases reliability data were collected after each tape placement.

Procedure

An ABAB reversal design was used to determine whether immediacy of supervisor feedback was functionally related to (a) rate of task completion and (b) the percentage of placement errors. The four phases of the study were Baseline$_1$, Immediate Feedback$_1$, Baseline$_2$, and Immediate Feedback$_2$. Both dependent variables were tracked across all four phases.

Baseline$_1$ During Baseline$_1$, supervisors provided feedback to the subject only after she raised her hand, signalling completed placement of all four tapes. Tapes were checked to determine if they met the quality control specifications. Supervisors praised the worker for correct tape placement and delivered one ounce of coffee, tea or soda if all four tapes were accurate. If a tape was placed incorrectly a supervisor told the subject what discrimination was in error ("too low", "not in the groove"). The supervisor then removed the incorrect tapes and returned the shield to the worker to be reworked. Only data on the initial completion of a shield was used to compute production rate. Supervisors' praise also declined as the subject worked for "on-task" behavior. This procedure ensured that the rate of staff contacts across all phases would be approximately the same. Baseline$_1$ procedures were in effect across the 20 days.

Immediate Feedback$_1$ During the 20 days of this phase, the subject was given feedback after each tape placement. Immediately after a tape was placed in the shield, a supervisor walked to the subject's work station and assessed the accuracy of placement. Correct placement was praised. If placement was incorrect, the supervisor removed the tape and gave it back to the subject. No explanation was given as to why placement was incorrect. The supervisor then stood behind the subject and provided physical assistance if necessary to assure accurate replacement of the tape. The supervisor left the work station as soon as placement was correct. Immediate feedback was provided for all tape placement, initial as well as replacements. When the subject completed an accurate shield she received praise and one ounce of coffee, tea or soda as in Baseline$_1$.

Baseline$_2$ Baseline$_2$ lasted 9 days and exactly replicated Baseline$_1$.

Immediate Feedback$_2$ This phase spanned 44 days and exactly replicated Immediate Feedback$_1$.

RESULTS

Placement of errors

Figure 1 presents the daily percent of placement errors across all phases of the study. During Baseline$_1$, the percent of tapes placed with one or more errors was gradually increasing. The subject's daily performance ranged from 16% errors to 100% errors during Baseline$_1$. Her placement errors during the last 5 days of Baseline$_1$ averaged an abysmal 68%. With the advent of Immediate Feedback$_1$ procedures there was a small drop in error rate to 55%. Over the course of the Immediate Feedback$_1$ phase the trend noted during Baseline$_1$ was reversed, with the subject peforming fewer and fewer errors. Her error rate during the last 7 days of Immediate Feedback$_1$ averaged 29%.

Baseline$_2$ data indicate another reverse in error trend with errors progressively increasing across the phase. The trend and variability during Baseline$_2$ substantially replicate the pattern observed during Baseline$_1$. The Baseline$_2$ increase in percent errors stabilized only after the Immediate Feedback$_2$ phase began. The subject's errors during Immediate Feedback$_2$ did not replicate the sudden drop noted when Immediate Feedback$_1$ was implemented.

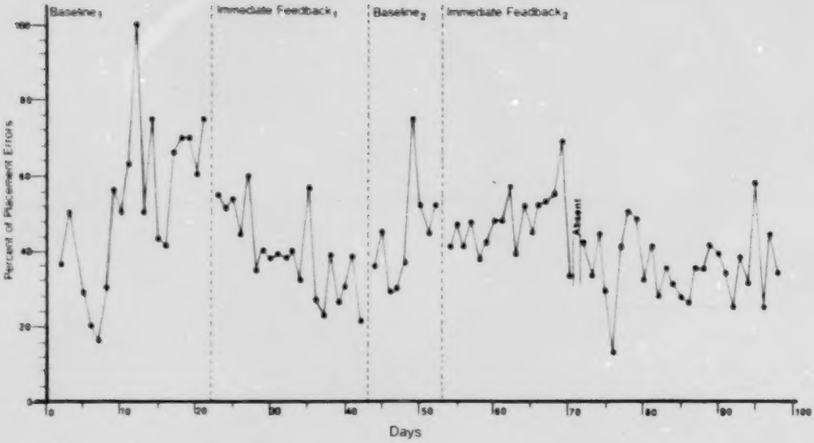


FIGURE 1

Percent of placement errors per day across phases

Over the 44 days of Immediate Feedback$_2$, however, there was a substantial reduction in the percent of errors. From an average error rate of 49.3% during the last three days of Baseline$_2$, the subject improved to an average of 35% errors over the last 28 days of Immediate Feedback$_2$.

Minutes per shield

Results for minutes per shield data are provided in Figure 2.

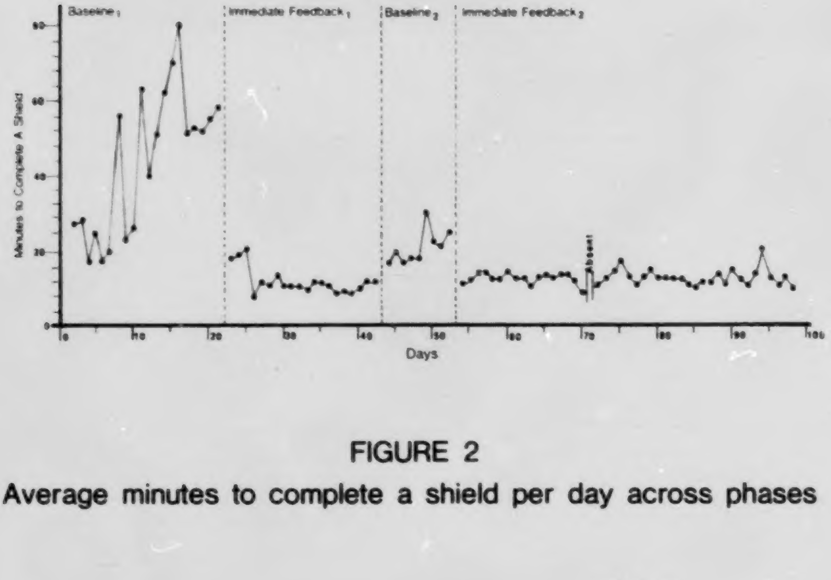


FIGURE 2

Average minutes to complete a shield per day across phases

During Baseline_1 there was a substantial increase in the average amount of time the subject took to complete a shield. During the first 6 days of the study the subject completed a shield on the average of one every 22.5 minutes. By the last 5 days of Baseline_1, however, the subject was averaging 53.8 minutes per shield. In addition to this deterioration in performance, there was substantial variability in response rate from day to day.

Following implementation of Immediate Feedback_1 procedures, there was an immediate drop in the minutes taken to complete a shield, and a major decrease in performance variability. During the last 17 days of this phase the subject averaged 10.8 minutes per shield within a stable, 6-minute range.

During the 9 days of Baseline_2 the subject began working progressively slower. Her average time to complete a shield increased from 12 minutes on the last day of Immediate Feedback, to 17 minutes on the first day of Baseline_2. This trend continued and the final three days of Baseline_2 averaged 22.7 minutes per shield. This was more than double the time required when immediate feedback was given.

The final phase of the study, Immediate Feedback_2, was characterized by an immediate reduction in the time taken for task completion, followed by a stable level of performance similar to that found in the Immediate Feedback_1 phase. Across the 44 days of Immediate Feedback_2 the subject averaged 12.5 minutes per unit. This is a sharp contrast from the gradually worsening performance noted in both Baseline phases.

DISCUSSION

School environments, work environments and in many cases the living environments of handicapped individuals require accurate performance of adaptive skills. The present study was designed to assess one easily-implemented method for improving the accuracy of work skills performed by a severely retarded woman. It was anticipated that differential feedback about the accuracy of tape placement would be more effective if the latency between the response (placing a tape) and the consequence (supervisor feedback) was minimized. It was further anticipated that improved accuracy would be correlated with improved rate of production.

The data indicate that the subject did perform with fewer errors, and at a faster speed during Immediate Feedback phases than during Baseline phases. From an experimental perspective, the ABAB reversal pattern for both dependent variables supports the existence of a functional relationship between the implementation of Immediate Feedback procedures and improvement in the subject's error and production performance.

These results are substantially tempered, however, by the fact that the subject's percent of errors at no time reduced to a level acceptable for industrial requirements. This limitation, coupled with the lack of a sudden drop in errors when the Immediate Feedback_2 phase was implemented suggest that while the

Immediate Feedback procedures were related to change in errors, this effect was not sufficiently powerful to produce a vocationally significant impact on the subject's errors.

The most powerful results from the present study are seen in the subject's dramatic improvement in production speed during Immediate Feedback phases. The level and trend of minute-to-completion data closely follow an ABAB reversal pattern across phases. In addition there is very little overlap between adjacent phases. This effect is particularly interesting given that supervisor's interruptions after each tape artificially increased the time taken to complete a shield.

While the ABAB reversal effect across phases is impressive the data do not suggest that this effect is the result of improved accuracy. Percent error data and minutes-per-shield data are not highly correlated during the Immediate Feedback phases. It is possible that while staff contacts were constant across all phases, the more intrusive praise and corrections delivered during Immediate Feedback phases functioned as more powerful reinforcing and/or punishing events than the praise delivered during Baseline phases. If this were the case the worker's improved performance may be simply the result of a shift in reinforcement schedules.

It has long been known that praise, attention and feedback affect the behavior of handicapped as well as non-handicapped individuals. As severely handicapped individuals become increasingly integrated in the mainstream of our society greater care and precision must be applied to determine just how these procedures can be best employed. The present study suggests that one factor for consideration may be the time delay between when a target response is performed and when feedback, or consequences, for the response are delivered.

References

Bellamy, G. T., Inman, D. P., & Horner, R. H. Design of vocational habilitation services for the severely retarded: The Specialized Training Program Model. In G. Hamerlynck (Ed.), Applied behavior analysis techniques for the developmentally disabled. New York: Brunner/Mazel, 1979.

Catania, C. A. Reinforcement schedules and psychophysical judgements: A study of some temporal properties of behavior. In W. N. Schoenfeld (Ed.), The theory of reinforcement schedules. New York: Appleton-Century-Crofts, 1970.

Federal Register, Chapter V, Part 525: Employment of Handicapped Clients in Sheltered Workshops. May 17, 1974, 17509.

Fink, W. T., & Carnine, D. W. Control of arithmetic errors using informational feedback and graphing. Journal of Applied Behavior Analysis, 1975, 8, 461.

Guralnick, M. J. Solving complex perceptual discrimination problems: Techniques for the development of problem-solving strategies. American Journal of Mental Deficiency, 1976, *81*, 18-25.

Jenkins, H. M. Sequential organization in schedule of reinforcement. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules*. New York: Appleton-Century-Crofts, 1970.

The Intra-Chain Response Patterns of Three Severely Retarded Adults During Extinction

Robert H. Horner, Bonnie Biel Wuerch, J. Daniel Boomer

Department of Special Education
University of Oregon

The authors express appreciation to G. Thomas Bellamy and Shawn M. Boles for their comments on earlier drafts of this paper.

This research was supported in part by a contract from the DHEW Bureau of Education for the Handicapped, Program Development Branch.

The Intra-Chain Response Patterns of Three Severely Retarded Adults During Extinction

Operant chains are ordered sequences of behavior with each step in the chain made up of a discriminative stimulus (S^d) and a response. Individual responses combine to form a chain when their accurate completion creates the S^d for the next response. In this way, the responses are linked together into an ordered sequence that typically terminates with delivery of a reinforcer. Operant chains have received considerable experimental analysis with respect to their acquisition, rate of performance and maintenance (D'Andrea, 1969; Ferster & Skinner, 1957; Gollub, 1977; Kelleher, 1966). A useful application of this research has focused on procedures for training and maintaining vocationally relevant operant chains with severely retarded adults (Bellamy, Horner & Inman, 1979; Crosson, 1966; Horner & Bellamy, 1978; White, 1970.)

A large number of vocational tasks available to handicapped persons require workers to perform operant chains. Accuracy, rate and maintenance of such chains are all variables that have received attention (Bellamy, Inman & Schwarz, 1978; Gold, 1973; Horner & Prill, in press; Martin & Pallota, 1979). It is becoming increasingly clear, however, that a comprehensive analysis of vocational chains requires study of variables that affect not only the entire chain, but intra-chain response patterns as well.

All responses within a chain do not change equally following an intervention. Characteristics such as response topography or position in the chain can affect specific intra-chain response patterns (D'Andrea, 1969; Gollub, 1977; Jwaideh, 1973). Horner and Bellamy (in press) have demonstrated that intra-chain stimulus control variables can affect the likelihood of certain irrelevant responses by severely retarded persons performing vocational tasks. Much needs yet to be learned, however, about the implications of intra-chain response patterns for programming strategies with handicapped persons.

Analysis of intra-chain responding can be viewed as analysis of stimulus control at each step in the chain. To date, a consistent finding in both the experimental and applied research is that stimulus control of chain steps is strongest towards the end of the chain (Gollub, 1977; Millenson, 1967). This effect is typically described with respect to delivery of the terminal reinforcer. As the proximity between a step response and the terminal reinforcer increases, there is an increase in the level of stimulus control exerted by the step S^d (Gollub, 1977). This relationship would predict specific intra-chain response patterns under different schedules of reinforcement. Extinction is one schedule for which such prediction is particularly appropriate. As stimulus control within an operant chain deteriorates under extinction, a more pronounced breakdown would be expected at steps early in the chain. While this pattern has proven predictive of intra-chain responding in experimental settings, (D'Andrea, 1969; Mackintosh, 1955) little effort has been directed at examining intra-chain patterns in applied settings. The present study examines the hypothesis that handicapped workers performing on an extinction schedule will work more slowly than when working on a continuous reinforcement schedule, and that this decreased rate will be a function of response deterioration early in the operant chain. Application of an operant chaining paradigm to vocational habilitation of retarded adults is examined, and implications for direct service procedures discussed.

Method

Subjects

Three adults labeled severely retarded on the basis of current administrations of the Stanford-Binet participated as subjects in the study. During the four years prior to the study, the subjects had been employed at the University of Oregon's Specialized Training Program, a subcontract workshop for severely retarded adults. Subjects A and B each could express basic needs via sign language and Subject C was capable of carrying on simple conversations. At the time of the study the ages of Subjects A, B, and C were 40, 24, and 52, respectively.

Task and Setting

The task used by all subjects required placement of eight 1 cm long, biaxle, electrical components in an 11 cm x 5 cm circuit board. The experimenter would place a printed circuit board blank in front of the subject and walk away. The subject was then required to lift the lid on a 10 cm x 18 cm plastic box, pick up one component, close the lid, and place the component in designated holes on the circuit board. This procedure was repeated until all eight components were in place. To maximize the homogeneity of the responses, all components were the same size, shape and color, and the position and orientation of holes on the circuit board were as similar as possible.

Subjects A and B placed components in the circuit board blank with their fingers. Subject C exhibited great difficulty performing this manipulation and as a result was trained to use a pair of needle-nose pliers to aid insertion. As a subject performed the task, his or her speed was evaluated for nine separate steps. The first step recorded the time between presentation of the circuit board blank and lifting the lid. The remaining steps included the time to pick up and place each of the eight components.

Apparatus

A standard Lehigh Valley data pac (LVE 421-09), power supply (LV 315-08) and multifunction timer (LVE 353-07) were used to record the number of seconds taken to place each component in the circuit board, and the total time to complete one circuit board. The printer was activated each time the subject opened the lid of the plastic box to remove a component. Termination of the final step, and activation of the printer was achieved via depressing a red key immediately adjacent to the plastic box. Throughout the study activation of the printer produced a noise audible to the subject.

Procedure

The study employed a single subject, multiple baseline design with an ABA reversal (Kratochwill, 1978). The dependent variables throughout the study were (a) average seconds per session to complete the entire chain (i.e. place all eight components in a circuit board), and (b) the average seconds per session to complete each of the nine steps of the chain. The independent variable was schedule of reinforcement. Two schedules were compared: continuous reinforcement for chain completion, and extinction.

Prior to the study reinforcer sampling was conducted for each subject. Subjects were given simultaneous access to five different consumables identified by their caregivers as high-preference items. An item was used throughout the study as an assumed reinforcer if it was chosen three consecutive times on each of three days. Based on the results of this procedure one-half ounce of coffee, cola or diet cola were used as contingent reinforcers for subjects A, B and C respectively.

In addition to reinforcer sampling each subject also received one-to-one training on the task immediately prior to the study. Training procedures closely followed those described by Bellamy, Horner and Inman (1979). Training was conducted during a 45 min training session with each subject each work day. The exit criteria from training was performance at 100% accuracy over two consecutive days with the trainer maintaining at least a 3 m distance from the subject. The first phase of the study began the day after a subject reached criterion. At this time the daily training sessions became experimental sessions. One experimenter conducted all sessions, with the exception of sessions occurring on days 20 through 24, which were conducted by a second experimenter. The study was composed of three phases: Baseline$_1$, Extinction, and Baseline$_2$.

Baseline 1: During Baseline$_1$ a subject entered the training room, and independently performed the task. The experimenter sat 3-5 m directly behind the subject. Upon completing the nine steps required for stuffing one circuit board a subject raised his or her hand. At this signal the experimenter approached the work table, praised the worker and delivered one-half ounce of the preferred liquid in a glass. When the subject set his or her glass back on the table the experimenter presented a new circuit board, reset the counters and immediately returned to his seat. No interaction occurred between a subject and the experimenter when a circuit board was presented, or while the subject was working.

Extinction: The Extinction phase replicated Baseline$_1$ only completion of a circuit board was followed by the experimenter approaching the subject, presenting another circuit board blank, resetting the printer and returning to his seat. No interaction occurred between the experimenter and the subject, and the identified reinforcer for each subject was not delivered.

Baseline 2: Procedures during Baseline$_2$ replicated Baseline$_1$ procedures. One-half ounce of the preferred liquid was again delivered each time a circuit board was completed.

Results

Overall Chain Performance

It was anticipated that data from the present study would indicate (a) that a subject's performance deteriorates during extinction, and (b) that this deterioration is most pronounced in the early portions of the operant chain. The results confirm both of these expectations. Figure 1 presents data on the average min per circuit board per day for each subject across all phases.

During the first 15 days of Baseline$_1$ Subject A performed at a very consistent, rapid pace, averaging approximately 1 min per circuit board. During this same period Subjects B and C demonstrated substantial improvements both in stability of performance and in rate of circuit board completion. Subject B dropped from a high of 6 min per circuit board to less than 2 min per board. Subject B improved from a high of 5.8 min per board to 3.2 min per board.

Following a multiple baseline format, Extinction was initially introduced only with Subject A. Soon after implementation of Extinction the subject's performance became increasingly variable, and his work speed decreased. As can be seen in Figure 1 the Extinction phase ended with Subject A averaging over twice as many sec to complete a board as he had during Baseline$_1$.

Analysis of Subject A's behavior change following the intervention must also focus on the performance of Subjects B and C at the time Subject A entered Extinction. Surprisingly, Subjects B and C demonstrated reversals in their improving trend on the day that Subject A entered Extinction. Subject B performed more slowly than she had in 12 days and Subject C worked an average

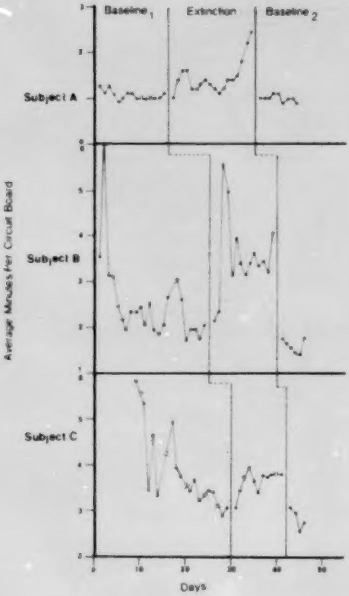


FIGURE 1

The Average Minutes to Complete A Circuit Board per Day across Phases for each Subject

of 1.5 min slower than she had the previous day. During later days of Baseline$_1$, both subjects returned to the improving trends they exhibited earlier.

Implementation of the Extinction phase with Subjects B and C was also followed by slower rates of performance. As can be seen in Figure 1 the increase in variability and increase in average min to complete a board observed with Subject A, were replicated by Subjects B and C.

With all three subjects initiation of Baseline$_2$ procedures was followed by an immediate improvement in performance. Subject A returned to his Baseline$_1$ pattern of fast, consistent working. Subjects B and C not only displayed dramatic improvements in speed and variability over their Extinction patterns but continued the improving trends noted during Baseline$_1$.

In general the data provided in Figure 1 indicate that implementation of Extinction was functionally related to deterioration of subject work speed. Experimental control provided by the multiple baseline design is slightly weakened by Subject B and C's brief reversals immediately following initiation of Extinction with Subject A. This potential threat to control is minimized, however, by the consistent performance patterns of all three subjects at all other phase change points, and by the use of an ABA reversal design with each subject. The

ABA design provides validation of the Extinction effect by adding within subject comparison to the already available across subject comparisons afforded by the multiple baseline format (Kratochwill, 1978).

Average Step Performance Patterns

While Figure 1 indicates the average time to complete a circuit board, Figure 2 shows the average time to complete each of the nine steps in a circuit board for each phase. The one consistent finding across all subjects is a major increase in the time taken to perform the first step under the Extinction phase. All three subjects displayed a small delay between presentation of a circuit board blank and lifting of the lid to begin working during Baseline conditions. In each case, however, this delay increased substantially during the Extinction phase; by a factor of 10 for Subject A, by a factor of 50 for Subject B, and by a factor of 15 for Subject C.

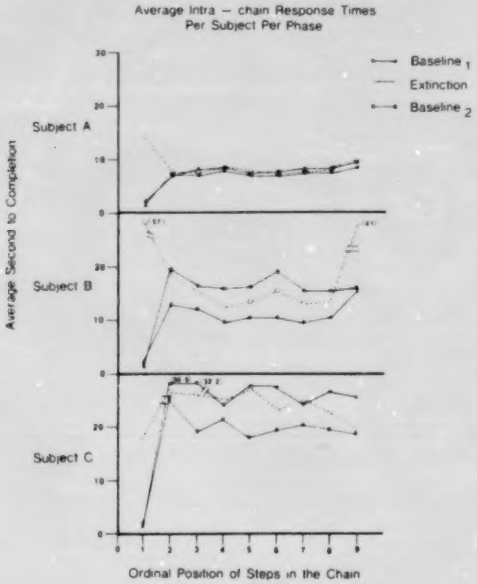


FIGURE 2

The Average Seconds to Complete each Step in the Chain across Phases for each Subject

As expected from Figure 1 data, Subjects B and C performed faster during Baseline$_2$ than they did during Baseline$_1$. This improvement appears related simply to performing all steps of the task at a faster pace. No specific intra-chain pattern discriminated between the two Baseline phases for either subject.

Subject B performed as anticipated during the Extinction phase, except that she took exceptionally long to perform the final step. This pattern may be a function of requiring subjects to press a red button as the final step in the chain. Reports from the experimenter indicate that Subject B quickly completed the responses of placing the components in the circuit board, but paused for several seconds before depressing the red button. This pause was not observed during either Baseline phase.

Subject A's Step Performance Pattern across Days

A more refined analysis of Subject A's intra-chain performance pattern in provided in Figure 3. This figure presents the average seconds to complete each step in the chain across all days. The data display a consistent, flat performance pattern during Baseline$_1$, and index the developmental changes that occurred across days during Extinction. The pattern of Subject A's data, though not the level or overall stability, is characteristic of data generated by Subjects B and C.

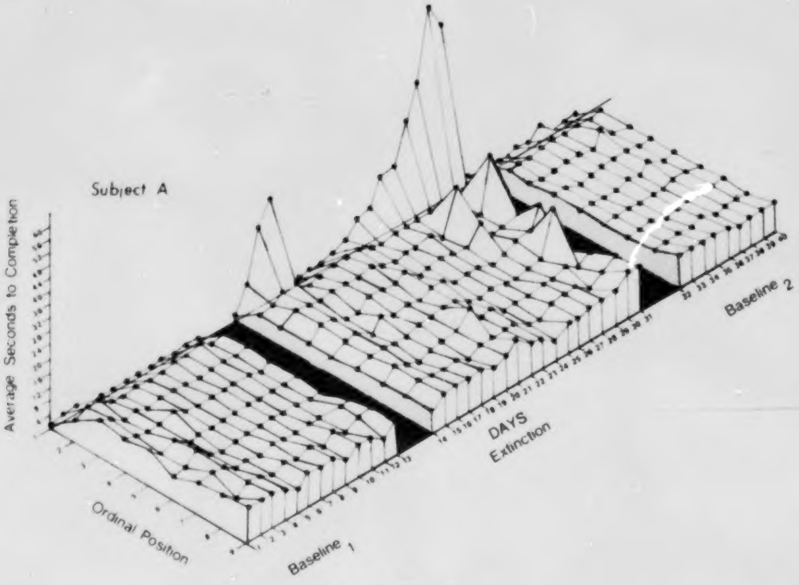


FIGURE 3

The Average Seconds to Complete each Step in the Chain across Days within each Phase for Subject A

Within the first 14 days of Extinction the only task step to show substantial change was step 1. The subject began taking more time before beginning work on the task. Once he began responding, however, the remaining steps were performed at a rate similar to that observed during Baseline$_1$. Only after 14 sessions of Extinction did deterioration begin at other steps of the task, and even these steps were located in the first half of the chain. An improvement in step 1 performance at the fifth day of the Extinction phase was correlated with the onset of a second experimenter conducting the experimental sessions.

This experimenter was present for five days during which Subject A again displayed a gradually increasing length of pausing before he performed the first response. A similar shift in performance did not occur for Subjects B and C. Subject A's trend for step 1 across days did not shift when the first experimenter returned. Reimplementation of baseline procedures was followed by an immediate return to Baseline_1 performance levels. Step 1 was performed without hesitation or pausing, and the intra-chain variability noted in the latter days of Extinction was replaced by smooth, rapid performance.

Discussion

The results presented in Figure 1 indicate that withdrawing reinforcement from severely retarded workers is related to reduced rates of productivity. The data on seconds to complete each step in the chain (Figure 2) demonstrate that this general deterioration in performance is most consistently related to pausing early in the chain. Close inspection of Figure 2 further indicates that, with the exception of step 1, Subjects B and C performed most steps faster during Extinction than they did during Baseline_1. With all three subjects, however, an overall slower work pace during Extinction was related to reduced performance on the first step in the chain.

Data provided in Figure 3 indicate that for Subject A the initial break-down in response rate occurred on the first step of the chain, but that in the later days of Extinction other steps began to show signs of deterioration. Even these steps, however, were early in the chain.

These data are consistent with predictions from the experimental literature, and support the assumption that even on a relatively short operant chain stimulus control is weaker at the beginning of the chain. The authors recognize, however, that rate of performance is but one of the response variables affected by extinction. Several researchers (Skinner, 1938; Millenson, 1967; Zeiler, 1978) have noted that in addition to affecting rate, extinction is related to (a) increased response variability, (b) changes in the topography of responding, and (c) increased intensity of responding. Data provided in Figures 2 and 3 support the expectation of increased variability during extinction. In addition Experimenter reports indicate that response topographies for all subjects changed after several days of extinction. Consistent reports document increases in pausing, orienting toward the Experimenter and self-stimulation during the Extinction phase. These observations suggest that subjects did not actually perform the task responses more slowly during extinction. Rather, when the responses were performed, the movements were at or near Baseline rates. During Extinction, however, new responses were added to the chain, and these new, irrelevant responses affected overall performance rates. No index of changes in response intensity were monitored during the study.

In general, data from the present study support the assumption that if a worker receives reinforcement only at the end of a chain stimulus control will be stronger toward the end of the chain than at the beginning. When the worker is placed on an extinction schedule the general rate of performance will decrease, and this decrease will be most pronounced in steps early in the chain.

These results hold substantial implications for managers of work environments employing severely handicapped persons. Hutton (1970) and Rolland (1969) have reported that the schedule for work completion in work environments employing handicapped persons is often near that defined by extinction. As such, high rates of work should not be expected until a consistent schedule of reinforcement is implemented. In addition, however, with severely handicapped workers it may be necessary to attend to intra-chain response rates as well as the rate of the entire chain. If a particular reinforcer or schedule of reinforcement becomes less effective, the first signs of this effect may appear in the early steps of the chain. Supervisory staff aware of this pattern may be able to intervene before work rate drops appreciably. In addition, it may be advisable for supervisors of severely handicapped workers to deliver praise or other reinforcers as workers are performing early steps in a task. This may function to strengthen responses that are likely to be under weak stimulus control.

If severely handicapped individuals are to gain competence in the individual living and vocational skills needed for successful participation in the mainstream of society, care will need to be focused on the programming details of service environments (Horner & Bellamy, 1979). The present study is an example of the level of detail that may be needed. The results indicate that even with a short, nine step task, variability in intra-chain responding can be identified and controlled. These results hold immediate implications for supervisory staff responsible for training and maintaining vocationally relevant operant chains. In addition they raise new questions such as what effects step difficulty, chain length, and delivery of intra-chain reinforcers will have on worker performance. In general the data provide support for the utility of experimental research in defining service issues in applied settings. In particular the data emphasize the value of applying an operant chaining perspective to the vocational habilitation of severely retarded adults.

References

Bellamy, G. T., Horner, R. H., Inman, D. P. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Bellamy, G. T., Inman, D. P., & Schwarz, R. Vocational training and production supervision: A review of habilitation techniques for the severely and profoundly retarded. In N. Haring and D. Bricker (Eds.), Teaching the severely and profoundly handicapped, Vol. 3. Columbus, Ohio: Special

Press, 1978.

Crosson, J. The experimental analysis of vocational behavior in severely retarded males. Doctoral dissertation, University of Oregon, Dissertation Abstracts International. 27:3304, 1966.

D'Andrea, T. Extinction of a heterogeneous chain after several reinforcement schedules. Journal of the Experimental Analysis of Behavior, 1969, 12, 127-135.

Ferster, C. B., & Skinner, B. F. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.

Gold, M. Research on the vocational habilitation of the retarded: The present, the future. In N. Ellis (Ed.), International Review of Research in Mental Retardation, Vol. 6. New York: Academic Press, 1973.

Gollub, L. Conditioned reinforcement: Schedule effects. In W. Honig and J. Staddon (Eds.), Handbook of operant behavior. New York: Prentice-Hall, 1977.

Horner, R. H., & Bellamy, G. T. A conceptual analysis of vocational training with the severely retarded. In M. Snell (Ed.), Systematic instruction of the moderately, severely and profoundly handicapped. Columbus, Ohio: Charles E. Merrill, 1978.

Horner, R. H., & Bellamy, G. T. Stimulus control within operant chains: Variables affecting intra-chain response patterns. In G. T. Bellamy and R. H. Horner (Eds.), Habilitation of severely and profoundly retarded adults, Volume III. Eugene, Oregon: University of Oregon, in press.

Horner, R. H., & Prill, N. The effects of immediate correction on the error rate of one severely retarded worker. In G. T. Bellamy and R. H. Horner (Eds.), Habilitation of severely and profoundly retarded adults, Volume III. Eugene, Oregon: University of Oregon, in press.

Hutton, S. A descriptive analysis of covariation of peer and/or teacher consequation of subject behavior with subject behavior. Master's Thesis, University of Oregon, 1970.

Jwaideh, A. R. Responding under chained and tandem fixed-ratio schedules. Journal of the Experimental Analysis of Behavior, 1973, 19, 259-267.

Kelleher, R. T. Chaining and conditioned reinforcement. In W. Honig (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966.

Kratochwill, T. R. Foundation of time-series research. In T. R. Kratochwill (Ed.), Single subject research: Strategies for evaluating change. New York: Academic Press, 1978.

Mackintosh, I. The resistance to extinction of responses acquired under irregular conditions of learning. Journal of Comparative and Physiological Psychology, 1955, 48, 363-370.

Martin, G., & Pallotta, A. Behavior modification in sheltered workshops and community group homes for the retarded: Current status and future considerations. In G. Hamerlynck (Ed.), Applied behavior analysis techniques for the developmentally disabled: Vol. 1. New York: Brunner Mazel, 1979.

Millenson, J. Principles of behavior analysis. New York: McMillan, 1967.

Rolland, J. An identification of person-environment social transactions and response patterns of trainable mentally retarded adolescents in different environmental settings. Doctoral dissertation, University of Oregon, 1969.

Skinner, B. F. The behavior of organisms. New York: Appleton-Century-Crofts, 1938.

White, O. R. An investigation of complex chains in the adult, retarded male. Unpublished master's thesis, University of Oregon, 1970.

Zeiler, M. D. Principles of behavior change. In A. C. Catania and T. A. Brigham (Eds.), Handbook of Applied Behavior Analysis. New York: Irvington Publishers, 1978.

PART III

Instrumentation For Program Evaluation

Multiple Perspective Evaluation of an In-Service Program for Vocational Trainers of Severely Retarded Adults

Nancy Weissman-Frisch
Fred A. Crowell
Dean P. Inman

Center on Human Development
University of Oregon

This work was supported by a grant funded jointly by the Department of Health, Education, and Welfare Developmental Disabilities Office, Region X and the Bureau of Education for the Handicapped.

The senior author would like to acknowledge Dean P. Inman for the initial STP in-service format, and Robert H. Horner for editing and reinforcement.

Multiple Perspective Evaluation of an In-Service Program for Vocational Trainers of Severely Retarded Adults

Recent federal legislation indicates a sincere national commitment to integrate severely and profoundly retarded adults into the mainstream of society (P. L. 94-103, P. L. 94-142, P. L. 93-112). Supporting the feasibility of this endeavor are numerous reports that severely and profoundly retarded adults can function independently in vocational settings given proper training and sufficient time (cf. Bellamy, Horner, & Inman, 1979). Ultimately, however, the success of current mainstreaming efforts is predicated heavily on having available sufficient numbers of highly skilled service delivery personnel to design and implement the necessary service programs. In-service programs are one way of efficiently meeting this need. Reliance on in-service programs for personnel preparation creates an expanding need for adequate evaluation instruments and procedures for assessing the effects of such training.

Recent publications reviewing issues in the evaluation of training programs (Grant & Anderson, 1977; Browning, 1974) indicate several inadequacies in current efforts to evaluate these programs. In general, little attention has been given to either program description or to the developmental, formative process that results in evaluative instruments:

> Limitations in available journal space undoubtedly contribute to inadequate description of the training programs. Nevertheless, such deficiencies leave the reader with little comprehension of what was evaluated. At the very least, *clear statements of the objectives of the program and a brief description of the methods employed in conducting the training would be helpful*. (Grant & Anderson, 1977, p. 67, italics added).

In addition, most evaluation designs necessitating instrument and program development, are strongly influenced by experimental research paradigms. These experimental research paradigms constrain both the trainer and the evaluator into positions of making no alterations in either the program or the instruments designed to evaluate the program (Edwards &

Guttentag, 1975; Mitroll & Bonoma, 1978; Patton, 1978).

In response to these inadequacies in both description and evaluation of training, the authors have focused upon four major dissemination objectives in this paper: (a) to briefly describe a multiple perspective evaluation approach and its application to an in-service seminar; (b) to describe an intense 4 1/2 day in-service training seminar designed to upgrade the service skills of vocational habilitation specialists; (c) to describe the multiple perspective evaluation package that has been developed to evaluate the impact of the in-service seminar; and (d) to discuss the implications of a multiple perspective evaluation approach to in-service training programs. A forthcoming paper will present the results accumulated through the use of the multiple perspective evaluation package.

Description of a Multiple Perspective Evaluation Approach

In contrast to single perspective evaluation, the authors have drawn upon a "multiple perspective" (Crowell, in press) approach to evaluating training programs. The multiple perspective allows two important processes to occur during instrument and program development: (a) the ultimate data users have important input as "participatory designers," in that they help develop the instruments and revise the program as data are accumulated, and (b) flexibility is achieved in permitting extensive changes to be made in both instruments and program components.

Multiple perspective evaluation also seeks to achieve a balance between the data needs of program design, program monitoring, and program development. Multiple instruments are essential to this approach in terms of both the variety of perspectives incorporated and the number of data generators and data users. The perspectives of data users are determined by their informational needs. Informational needs, in turn, are generated from the problems of the data users. Problems, in this context, are defined as discrepancies between perceived conditions and preferred conditions. One major distinction between this approach and single perspective evaluation approaches lies in the amount of involvement program data generators and users have in the total evaluation process, as "discrepancy" definers.

In the present evaluation contexts, as in most in-service contexts, these data users and data generators with their distinct perspectives and informational needs are represented by the seminar instructors, the seminar participants, the program director, the program evaluator, and the in-service coordinator. From the seminar instructor's perspective, data are required related to presentation of in-service content and the attainment of specific in-service objectives. In addition, instructors need to receive formative reports summarizing in-service data derived from a number of instruments on a regular basis. The participants, as professionals who train and supervise retarded persons, are interested in data related to how the in-service meets their objectives. The program director requires data that are more "summative" and product oriented that would be of use to potential consumers of the in-service training package. The in-service coordinator and the program evaluator are interested in how many participants achieve what kinds of competencies. The coordinator and evaluator are also sensitive to the differences between perspectives of instructors, participants, and the program director. They generate and use data that detect and communicate these differences as complimentary representations of the training process.

Description of the In-Service Seminar

The in-service seminar at the Specialized Training Program (STP) was designed to provide skills to individuals who work with severely and profoundly retarded adolescents and adults (see Table 1 for description of program participants) in the areas of task analysis, training, and production supervision. The seminar is held once per month, beginning at 8:00 a.m. Monday morning and continuing through Friday noon. Five to seven individuals attend each seminar. The three major areas of task analysis, training, and production supervision are broken down into smaller instructional units which are presented sequentially. (The content of these units is described in greater detail in Bellamy, Horner, & Inman, 1979.) Each unit is designed to train

TABLE 1
Descriptive Characteristics of 1977 - 1978 STP In-Service Training Seminar Participants and Their Clients

Descriptor	Number of Participants/ Mean Years
Total Number	57
Mean Age	29
Sex: Female	30
Male	27
Geographical Distribution	
Northwest	45
Elsewhere (U.S.)	12
Job Titles	
State Program Coordinators	5
Facility Director	4
Program Coordinators	4
Special Education Teachers	8
Workshop/Activity Center Staff	32
Other	5
Average Job Tenure	2.5
Average Job Experience	8.3
Type of Client Disability	
Retarded	42
Multiple Handicapped	13
Emotionally Disturbed	1
Culturally Disadvantaged	1
Age Level of Clients	
Adults	47
All Ages	2
Adolescents and Adults	2
Children and Adolescents	4
Adolescents	1
Children	1

program participants to perform specific habilitation skills at or above a minimum level of competence. After each skill is modeled by instructors, participants are provided an opportunity to practice the skill being taught.

There has been little research to date on methods of teaching behavioral techniques to workshop and activity center personnel who work with severely and profoundly retarded adults. Literature is available, however, describing competency-based programs for teachers of the severely and profoundly handicapped (Wilcox, 1977; Horner, 1977; Burke & Cohen, 1977), and on the efficacy of training institutional attendants, group home staff, parents, and teachers of the retarded in the use of behavior modification techniques (Gardner, 1972; Schinke & Wong, 1977; Panyan & Patterson, 1974). Modeling (Panyan & Patterson, 1974), role-playing (Gardner, 1972), and instructor feedback (Parsonson, Baer, & Baer, 1974) have been demonstrated to be useful techniques in training these groups. These techniques are utilized in the in-service described in this report. For an overview of in-service content see Table 2.

TABLE 2
In-Service Content

Schedule of In-Service Units	Method of Instruction	Unit Outcomes
I. Day One A. Pretests		
B. Writing Behavioral Objectives	Didactic, modeling, practice with feedback	Final behavioral objective which includes: 1) condition under which behavior is to be performed, 2) the behavior, 3) the criterion to which the behavior is to be performed
C. Behavioral Observation	Didactic, modeling, practice with feedback (observation in STP Work Activity Center)	Reliability Index of at least 85%
D. Task Analysis	Didactic, modeling, practice with feedback	Final task analysis which includes: 1) sequenced steps, 2) cues for each step, 3) criterion for each step
II. Day Two A. Introduction to Training	Didactic, video taped training models	Worksheets on training procedures
B. Difficult Steps Training	Didactic, modeling	Written identification of difficult discriminations and manipulations/ and remediating training strategies
C. Simulated Training	Modeling, practice with feedback	Training another participant on a task using four specific training techniques
III. Day Three and Four A. Training (3 participants each day)	Modeling, practice with feedback	Use of specified training techniques at an accuracy level of at least 85%
B. Production Planning (3 participants each day)	Didactic, observe shop	Written identification of two instances of production planning in STP shop
C. Production Supervision	Modeling, practice with feedback	Use of specified production supervision techniques at an accuracy level of at least 85%
IV. Day Five A. Ethics	Didactic	
B. Film		
C. Posttests		
D. Feedback to participants Feedback from participants	Discussion	

Application of Multiple Perspective Evaluation: An Evaluation Package

The purpose of applying multiple perspective evaluation to the in-service seminar was to allow all persons involved with the seminar to participate in the development and modification of instruments, and to receive data relevant to their individual data needs. This application has resulted in an evaluation package. The package includes a variety of instruments and formats in which to present the data collected with these instruments. In the following sections, instruments and data presentation formats are presented in the context of relevant data users. For an overview of evaluation instruments, data formats, and data users, see Table 3.

TABLE 3

Overview of Evaluation Instruments, Data Formats and Data Users

Program Components	**Program Data Users**			
	Seminar Instructors	Seminar Participants	Program Director	In-Service Coordinator-Evaluator
Training	-TEK -TDA -GAI	-TEK	-Telephone Interview -Attitude Survey	-Multiple Choice Concepts -Training Skills
Task Analysis	-TEK -TDA -GAI	-TEK	-Telephone Interview -Attitude Survey	-Multiple Choice Concepts -Task Analysis Skills
Production Supervision	-TEK -TDA -GAI	-TEK	-Telephone Interview -Attitude Survey	-Multiple Choice Concepts -Observation Skills -Production Planning Skills -Objective Writing Skills -Supervision Skills
Program Evaluation	Formative Reports		Summative Reports	-Meta-systemic Data

Seminar Instructors

Instructors conduct ongoing evaluation of the seminar using three instruments: The Training Evaluation Kit Program booklets (TEK, Browning & Foss, 1977), a Trainer Decision Analysis (TDA) constructed by the project, and with a Goal Attainment Index (GAI) constructed by the project. In-service instructors are also data users of formative reports. A description of these instruments and their uses is provided in this section.

Training Evaluation Kit (TEK). In-service participants are instructed in the TEK booklet to rate each in-service unit at its completion. The material presented, the presentation, and the presentor are rated on a four point scale. For reliability and validity data see Browning and Foss (1977). These ratings provide in-service instructors with an evaluation of each of their units on a monthly basis. Instructors can modify their presentations based on ratings.

Trainer Decision Analysis (TDA). The TDA is a form filled out by instructors one week subsequent to each in-service. It allows instructors to weight and rate the attainment of instructor-defined, general in-service objectives in the three in-service areas of task analysis, training, and production supervision. Instructors are presented with six objectives. They are asked to weight each objective, so that a total of 100 points is distributed among the objectives. They rate attainment of objectives in their area on a 1-9 scale. The TDA instrument provides instructors with the opportunity to revise expectations and priorities related to each component in their weightings and ratings. In addition, they can evaluate their scores in the context of the scores assigned by other instructors.

The Goal Attainment Index (GAI). The GAI, derived from a similar instrument (Fast, 1974), instructs participants to weight and rate specific in-service objectives for each in-service component. Participants are presented with 13 in-service objectives in the areas of task analysis, production supervision, and training at the beginning of the in-service week. They are instructed to weight each objective so that a total of 100 points is distributed among the objectives in each area. At the end of the in-service week, participants are instructed to rate the attainment of these objectives. This provides instructors with an index of how well participants see specific in-service objectives as having been met.

Formative reports. Formative reports occur on a monthly basis. These reports are a summarization of the previous month's data from all of the in-service instruments. They allow instructors to receive an interpretation of all the in-service data by component and instrument. Formative reports are disseminated in a monthly meeting, and are posted with graphs in a convenient place for users to view. Instructors may use these data to evaluate how well seminar participants have performed in all instrument-related areas.

The Seminar Participants

Data collected in portions of TEK are used by seminar participants. These portions include ratings of participant objectives, ratings of participant changes in four specified areas, and ratings of actual and preferred time spent in specified in-service areas. This section of the report provides a description of these data formats and data uses.

Training Evaluation Kit. In one portion of TEK, participants are instructed to write their objectives for the in-service.

This occurs at the beginning of the in-service week. Participants are then instructed to rate the attainment of these objectives at the end of the in-service week. Participants are also instructed to rate changes in themselves in the defined areas of behavior, motivation, information, and attitude in the TEK booklet. They also rate preferred and actual time spent in these four areas. These measures provide participants with an opportunity to define their own in-service objectives, an index of the usefulness of the in-service in meeting these objectives, an index of change, and a format to evaluate time spent during the in-service. Validity and reliability data is provided by the TEK user's manual (Browning & Foss, 1977).

The Program Director

Data for the program director of the in-service are collected with the telephone interview and the attitude survey. Summative data reports are also of interest to the program director. The data formats and data uses are as follows:

Telephone interview. The telephone interview was developed to assess how the participant rates his/her performance in the areas of task analysis, training, and production supervision after ascertaining the amount of time s/he spends in each of those areas. The interview was developed by the in-service coordinator and the program director. This portion of the interview is given once in written form on Monday of the in-service week and a second time over the phone 10-12 weeks later. A post-only section is also given over the phone inquiring into in-service techniques utilized at work, the number of staff trained in techniques, and the number of workers affected by this usage. In addition, participants are asked to rate their job satisfaction on a 1-5 scale before and after the seminar. Another section of the telephone interview provides for evaluation of participant objectives in their work environments. Participants are instructed to write objectives for themselves in their work environments at the end of the in-service week. They rate their attainment of these objectives on a 1-9 scale during the follow-up phone call. This information provides the program director with data related to the effect of the in-service on participants and their colleagues.

Attitude survey. An attitude survey was added in the middle of the in-service year. It was designed to assess whether or not participants' comments on changes in their attitudes towards the vocational skills of severely and profoundly retarded adults were in fact accurate and/or measurable. Studies have been previously conducted to assess changes in attitude towards retarded persons as a result of learning materials (Prothero & Ehlers, 1974) and tours (Sellin & Mulchahay, 1965).

STP staff members were asked to list what they considered to be important vocational and social attitude areas concerning severely and profoundly retarded adults. Changes, if they occurred to in-service participants, were expected in the area of attitudes towards the vocational skills of severely and profoundly retarded adults. Questions about mildly retarded adults and about social behaviors were included for comparative purposes (Hollinger & Jones, 1970).

A mapping sentence (Jordan, 1971; Castro & Jordan, 1977) was devised, and all items were written in terms of the mapping sentence (see Figure 1). The mapping sentence, much like an "item frame" in criterion-reference measurement, allows a data user to specify the various facets of a domain of items and to generate all possible combinations of facet values as instrument items. The use of competent trainers in the generation of objectives, which were then used to identify facets, insured content or descriptive validity. Such item validity was increased by having the same training staff members respond to each item generated with the mapping sentence, and including only those items with 90% agreement on the extreme item values (1, 2, 5, 6). A decision consistency score of 5% (absolute difference) was obtained on the same group of trainees with one month separating the two administrations both prior to exposure to in-service training activities. This information allows the program director to assess the effect of the in-service on participants' attitudes towards the vocational skills of severely and profoundly retarded adults on a validated instrument.

Subject (A)		**Referent** (B)	
myself a_1		mildly retarded adults b_1	
according to	believe that	severely and profoundly retarded adults b_2	
my group a_2			
Valuation (C)	**Referent Behavior** (D)	**Task Descriptor** (E)	
are able C_1	learn d_1	simple e_1	
should be able C_2	perform d_2	complex e_2	tasks
	have d_3		
Setting Descriptor (F)		**Condition** (G)	
social f_1		opportunities g_1	
in vocational f_2	settings given	training g_2	
legal f_3		normal conditions g_3	

FIGURE 1

Mapping Sentence Used to Generate Attitude Survey Items

Summative report. The summative report occurs at the end of the in-service year. It is of interest to the program director as it is a discussion of all of the preceeding year's data accompanied by graphs and tables. It provides the program director with data related to skills and information acquired and utilized by in-service participants. It is generally presented in written and verbal forms. Portions of the summative report may be included into federal and state grant reports.

The In-Service Coordinator and the Program Evaluator

Data used by the in-service coordinator and the program evaluator are collected with a number of instruments. Included are performance instruments, the multiple choice test, and all other instruments in a "metasystemic" fashion. Data formats and data uses will be provided in this section.

Performance instruments. Performance instruments are those instruments that measure participants behavior. Each behavior is related to an in-service unit objective. These instruments were developed and modified as a result of input from seminar instructors, the STP staff, the in-service coordinator, the program evaluator, and in-service participants.

Participants' attainment of the following unit is evaluated during each in-service unit. Instructors write participants' raw scores where appropriate, and circle a "0" or a "1" on participants' scoring sheets when evaluating participants (see Figure 2). A "0" indicates unsuccessful completion of the unit, and a "1" indicates successful completion. Participants are expected to receive a "1" for each unit prior to completing the in-service. Participants are expected to attain a "3" out of "3"

Name ______________________ **Date** ____________

Day 1	Identify three components of a behavioral objective	Write a behavioral objective	Reliably monitor problem behaviors	Identify S, R, & criterion	Sequence a task	Write a task analysis
	0 or 1	0 - 3 0 or 1	90% 0 or 1	0 or 1	0 or 1	0 - 6 0 or 1

Day 2	Define rules & procedures for training	Identify two ways to minimize difficult discriminations	Simulated Training
	0 or 1	0 - 4 0 or 1	0 or 1
Day 3	Identify two instances of production planning	Define and implement supervision contingencies	Implement Training Procedures
	0 or 1	85% 0 or 1	85% 0 or 1

FIGURE 2
Participant Scoring Sheet

on a final written behavioral objective for inclusion of all three behavioral objective components, and a "5" out of "6" on a final written task analysis for inclusion of task analysis components. A "2" out of "2" is expected on written identification of training techniques while viewing the training video tape, and a "4" out of "4" on written identification and ways to train difficult discriminations and manipulations in a given task. All written evaluation materials are scored by two instructors familiar with that in-service unit.

Attainment of an 85% inter-observer agreement is expected between participants, recorded observations of worker behavior in the behavioral observation unit. This inter-observer agreement is computed by totaling agreements across each 10-second observation interval and converting the total to a percentage $\left(\frac{\text{Agreements}}{\text{Agreements \& Disagreements}} \times 100\right)$. An agreement occurs when both participant observers note the occurrence or non-occurrence of a defined behavior during the same interval.

In the production supervision and training units participants are expected to attain a score of at least 85% on the appropriate use of techniques. Instructors observe and record participants' behavior on data collection forms designed in accordance with unit objectives. Instructors record the occurrence or non-occurrence of defined training and production supervision techniques. During training, data are collected on participants' training responses every ten seconds, in the following categories: 1) interrupt, 2) reinforce, 3) observe, 4) physical assist, 5) verbal cue, 6) physical cue, and 7) model. Observation periods are 15 minutes in duration. Observations of participants' use of production supervision procedures occur continuously across the following categories: 1) reinforce individual workers at specified variable interval schedules, 2) verbally, physically, and tangibly reinforce individual workers for on-task and task completion behaviors; 3) record production and interaction data; 4) respond to appropriate worker hand raises at task completion; 5) perform quality control checks at task completion; and 6, 7, 8, 9) carry out four specified behavioral programs with four workers. Data are also collected on participants' use of a scanning technique at 30-second intervals. Observation periods are 10 minutes in duration. Observer-instructors are trained, on video tapes and in vivo, to an inter-observer agreement level of 85%. Reliability checks are taken on instructors periodically to insure accurate data taking. Performance instruments provide the in-service coordinator and the program evaluator with data on the attainment of in-service criteria by participants.

Multiple choice test. The multiple choice test was incorporated to assess participants' knowledge of terminology and techniques used during the in-service. The pretest is given on day one and the posttest on day five of each in-service week. Items were generated for each of the components of task analysis, training, and production supervision. All in-service instructors generated multiple choice items. Items were eliminated if participants and staff persons, who were asked to take tests, consistently answered the questions incorrectly. In-service participants were also given time to discuss questions and answers with the in-service coordinator. Those comments were recorded and questions were modified accordingly. Two experts in the habilitation field were then given the items and component objectives and instructed to judge each item on congruence (+), incongruence (-), or undecided (0). The resulting data were analyzed to yield an index of item-objective congruence (Rovinelli & Ambleton, 1976) for each item. Only items with an index of .75 or higher were retained. The total item pool for each component domain was then randomly divided into two subtests and both tests (intermingled in one form) were administered twice to a group of direct care persons with approximately a week separating the two administrations. The results were analyzed to yield a mean absolute difference score for each of the three components as shown in Table 2. The closer the percentage score is to zero, the more reliable the test is for the population being examined (Millman, 1974).

The decision consistency of each component test was also checked in terms of percentage of mastery and non-mastery decisions. As can be seen in Table 4, the results again indicated by absolute difference between first and second test administration, a high decision·consistency for all three components. The data analysis also reveals that the task analysis items were more difficult than the other two, but that the alternate forms were of equal difficulty for all three component tests. The final forms consist of 28 questions each.

TABLE 4

Multiple Choice Reliability Data: Decision Consistency Percentages for Training, Task Analysis, and Production Supervision Components

Program Component	Mean Absolute Difference Score (%)	Mastery Decisions (Absolute Difference %)	Mean Errors	
			1st	2nd
Training (20 items)	3.3	4.6	4.3	3.8
Task Analysis (16 items)	10.5	4.6	6.7	6.1
Production Supervision (20 items)	7.0	4.5	5.1	5.4

Data collected with this instrument provide the in-service coordinator and the program evaluator with an evaluation of participants' knowledge of in-service information. Results from pretest and posttest multiple choice items allow for an assessment of participants' knowledge prior to and subsequent to the in-service. Data about the test itself provide the evaluator with an index of the tests' validity and reliability.

Meta-systemic data. The primary data generated and used by the in-service coordinator and the program evaluator is "meta-systemic" (Beer, 1970; Crowell, 1979). A "meta-system" is any system of a higher logical type than some referent, objective system of interest. Meta-systemic data provides the coordinator and evaluator with a description of the training system and the differences between perspectives of data generators and users in that system. Such descriptions cannot be generated from any single instrument or data source. Multiple instruments and data sources are necessary to provide this process of system self-description. For example, participants view the in-service seminar as a means of solving training and production problems within their organizational settings. Their perspective and the data that support it are complimentary to the perspective of the STP program staff who view the seminar as a means of providing training services to direct care staff and disseminating the results of research and development activities in vocational training for severely retarded adults. The coordinator and evaluator are able to generate information based upon the differences between these perspectives and share the information with all the data users concerned. These users then have a more valid and complete representation of actual in-service processes and products in relation to preferred processes and products.

Implications of a Multiple Perspective Approach to Evaluating In-Service Training Programs

In this report the authors have described an in-service seminar for trainers of handicapped persons and the evaluation package designed to provide data for multiple users. The multiple perspective evaluation approach addressed the data needs of instructors, participants, the program director, the coordinator, and the program evaluator. Instruments were developed and revised as each data user provided feedback to the coordinator and evaluator concerning the adequacy of the instruments in meeting his or her needs. The overall result was an evolving in-service program with a flexible evaluation design.

Traditional evaluations of in-service programs are usually conducted within a single perspective context, typically the research context. That is, evaluation problems have been conceptualized as "research" or "model-building" problems rather than "decision" or "political" problems (Popham, 1975; Cronbach, 1977). The major consequence of this single perspective approach is a methodological rigidity that prevents program evolution and "participatory design." Only one group of data users are served at the expense of other users. These other users' data needs are rarely identified and generally remain unmet.

There are, however, a number of problems associated with this approach which should be brought to the attention of service staff seeking help in designing and evaluating training programs. One problem is the reluctance of some staff to be flexible and tolerate change in either instruments or program components. Other problem areas include the higher costs of multiple instruments and the potential for "data overload." In addition, methods of summarizing data and presenting data in a format that is useful to different data users pose problems that are usually not associated with single perspective evaluation. The use of the Trainer Decision Analysis (TDA) instrument, the post-seminar meeting for discussing and interpreting data, and the posting of monthly summaries and graphs, are attempts to address these problems. Such meetings encourage instructors, program directors, and in-service coordinators to use data in formative, change-inducing ways for both programs and instruments. If data are not useful in making program decisions or are not being displayed in ways that communicate to these users, then corrective action cannot be taken to modify programs, instruments, and/or displays. Program design and program evaluation thus become a shared responsibility with the blending of multiple perspectives.

In spite of higher development costs and other "change problems," the benefits of the present approach to trainers, trainees and their organizations can be summarized as: (1) an evaluation that is more responsive to the informational needs of data users; (2) a more comprehensive and effective training experience; (3) a training program that evolves to accommodate the multiple inputs from trainers, trainees and program staff.

Although the focus in this report has been upon describing the application of a multiple perspective evaluation approach to an in-service seminar, the importance of a more detailed seminar content and actual data supportive of such content should not be neglected. Accordingly, a forthcoming report will focus upon the performance of seminar participants who have received training and are using that training in some direct service capacity.

References

Beer, S. Managing modern complexity. Futures, 1970, 2, 114-122.

Bellamy, G. T., Horner, R. H., & Inman, D. P. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Browning, P. L. Evaluation of short-term training in rehabilitation, Vol. 3. Eugene: University of Oregon, Research and Training Center, 1974.

Browning, P. L., & Foss, G. Training evaluation kit (TEK). Eugene, Oregon: University of Oregon, Research and Training Center, 1977.

Burke, P. J., & Cohen, M. The quest for competence in serving the severely/profoundly handicapped: A critical analysis of personnel preparation programs. In E. Sontag (Ed.), Educational programming for the severely and profoundly handicapped. Division on Mental Retardation: The Council for Exceptional Children, 1977.

Castro, J. G., & Jordon, J. E. Facet theory attitude research. Educational Researcher, 1977, 6(11), 7-11.

Cronbach, L. J. Remarks to the new society. Evaluation Research Society Newsletter, 1977, 1(1).

Crowell, F. A. Evaluating programs for educating mentally retarded persons: Changing paradigms. In J. Gottlieb (Ed.), Educating the mentally retarded person. Baltimore, Maryland: University Park Press, in press.

Edwards, W., & Guttentag, M. Experiments and evaluations: A re-examination. In C. Bennett and A. Humsdaine (Eds.), Evaluation and experiment: Some critical issues in assessing social programs. New York: Academic Press, 1975.

Fast, D. A new approach to quantifying training program effectiveness. Training and Development Journal, 1974, 28, 8-14.

Gardner, J. N. Teaching behavior modification to non-professionals. Journal of Applied Behavioral Analysis, 1972, 5, 517-521.

Grant, D. L. & Anderson, S. B. Issues in the evaluation of training. Professional Psychology, 1977, 8, 659-673.

Horner, R. D. A competency-based approach to preparing teachers of the severely and profoundly handicapped: Perspective II. In E. Sontag (Ed.), Educational programming for the severely and profoundly handicapped. Division on Mental Retardation: The Council for Exceptional Children, 1977.

Hollinger, C. S., & Jones, R. L. Community attitudes toward slow learners and mental retardates: "What's in a name?" Mental Retardation, 1970, 8, 19-23.

Jordan, J. E. Construction of a Guttman facet designed cross-cultural attitude-behavioral scale toward mental retardation. American Journal of Mental Deficiency, 1971, 76, 20-219.

Millman, J. Criterion-referenced measurement. In W. J. Popham (Ed.), Evaluation in education: Current applications. Berkeley: McCutchen, 1974.

Mitroff, I. I., & Bonoma, T. V. Psychological assumptions, experimentation, and real world problems. Evaluation Quarterly, 1978, 2, 235-260.

Panyan, M. C., & Patterson, E. T. Teaching attendants the applied aspects of behavior modification. Mental Retardation, 1974, 12(5), 30-32.

Parsonson, B. S., Baer, A. M., & Baer, D. M. The application of generalized correct social contingencies: An evaluation of a training program. Journal of Applied Behavior Analysis, 1974, 7, 427-437.

Patton, M. Q. Utilization-focused evaluation. Beverly Hills: Sage, 1978.

Popham, W. J. Educational evaluations. Englewood Cliffs: Prentice-Hall, 1975.

Prothero, J. C., & Ehlers, A. H. Social work students' attitude and knowledge changes following study of programmed materials. American Journal of Mental Deficiency, 1974, 79(1), 83-86.

Rovinelli, R., & Ambleton, R. K. On the use of content specialists in the assessment of criterion-returned test item validity. Laboratory of Psychometrics and Evaluative Research, Report No. 24. Amherst, Mass: The University of Massachusetts, 1976.

Schinke, S. P., & Wong, S. E. Evaluation of staff training in group homes for retarded persons. American Journal of Mental Deficiency, 1977, 82(2), 130-136.

Sellin, D., & Mulchahay, R. The relationship of an institutional tour upon opinions about mental retardation. American Journal of Mental Deficiency, 1965, 70, 403-412.

Wilcox, B. A competency-based approach to preparing teachers of the severely and profoundly handicapped: Perspective I. In E. Sontag (Ed.), Educational programming for the severely and profoundly handicapped, Division on Mental Retardation: The Council for Exceptional Children, 1977.

A Telephone Technique for Assessing Living Environments and Lifestyles of Retarded Adults

Nancy Weissman-Frisch, Fred A. Crowell,

G. Thomas Bellamy, and David Bostwick

Preparation of this chapter was supported in part by a grant from the Bureau of Education for the Handicapped, Research Grant #G00-760-5455.

A Telephone Technique for Assessing Living Environments and Lifestyles of Retarded Adults

Introduction

Ideological, economic, and programming factors have converged during the last decade to support dispersion of services for retarded people throughout normal community settings. One result of this dispersion is that program planners, funders, and advocates have restricted access to information about the nature and effectiveness of available services.

Information on the lifestyles of retarded individuals seems particularly important for decisions about service programs. To date, efforts to obtain such data have involved naturalistic observation systems (Landesman-Dwyer, Stein, & Sackett, 1978), ethnographic techniques (Edgerton, 1975; Edgerton & Langness, 1978) and mailed questionnaires and rating instruments (Bell, 1976; Gollay, 1977; O'Connor, 1976). While the utility of each of these methods of lifestyle description has been clearly demonstrated, none has been used extensively for ongoing program monitoring. The cost of observation and ethnographic techniques and the logistical difficulties typically encountered with mailed questionnaires reduce the utility of these approaches for repeated program evaluation efforts. To generate regular evaluative data, therefore, less expensive instruments are needed that can be used to assess service quality and individual lifestyles in community programs.

The Telephone Interview on Living/Leisure Environments (TILE) was conceptualized and designed within this problem context and as a part of a multiple assessment strategy (Irvin, Crowell & Bellamy, 1979) for evaluating a vocational training program for severely retarded adults (Bellamy, Horner & Inman, 1979). TILE applies the structured telephone interview methodology developed by Jones (1974) to obtain data on selected activities of retarded individuals.

This report provides an overview of TILE and presents a descriptive study that illustrates the information provided by the instrument and evaluates its methodological characteristics.

Instrument Description

Administration of TILE involves a structured telephone conversation of approximately 15 minutes in which an interviewer who is familiar with the instrument obtains information about a person from a caregiver who has direct contact with that person. The interviewer reads questions in order from an interview form, and records the respondent's answers on the TILE forms. Calls are repeated at weekly intervals at preselected times for four weeks.

The interview form is divided into three sections. The first section, an information background sheet, allows the interviewer to obtain descriptive data on the retarded subject, the respondent, and the residence. The second section is a checklist of 11 behaviors. Caregivers are asked if the subject performed the behaviors once, more than once, or not at all during the previous 24 hours. The third section of the interview form lists 16 behaviors and program events that are rated by the caregiver as occurring once, more than once, or not at all during the previous week. Two final questions in the form allow the respondent to rate (on a 1-5 scale) his or her overall satisfaction with the subject's behavior, and to report whether anything unusual happened during the last week. The Interview Form is presented in Table 1.

TABLE 1

TILE Interview Form

TILE INTERVIEW FORM

Interviewer ______________ **Interviewee** ______________

Date ______________ Resident ______________

2nd Interviewee (reliability) ______________

I. Did ______________ do any of the following during the last 24 hours?

Activity	Yes	If yes, once	more than once	elaborate
1. Argue				
2. Complain				
3. Initiate social interaction				
4. Display affection in a socially acceptable way				
5. Watch TV for at least 2 hours				
6. Engage in stereotyped repetitive behaviors				
7. Cry				
8. Dress independently				
9. Require help washing				
10. Require help toileting				
11. Complete a regular chore without a reminder				

CONTINUED ON NEXT PAGE

II. Did ____________ do any of the following during the last week?

1.	Steal				
2.	Fix (repair) something				
3.	Destroy property accidentally				
4.	Use tools independently				
5.	Leave home w/o permission				
6.	Fight				
7.	Leave home for an activity other than work a. community activity				
	b. visit family				
	c. visit friend same sex				
	opposite sex				
	d. go out to meals				
	e. go shopping				
	f. go for medical treatment				
	g. field trip				
	h. other				
8.	Destroy property deliberately				
9.	Tantrum				
10.	Was there a turnover in staff who have contact w/____________?				
11.	Was ________ the subject of a group home staff meeting?				
12.	Was ________ the subject of a meeting of group home staff and staff outside of the group home?				
13.	Did ________ receive medical attention from a physician or nurse?				
14.	Was there a change or problem in funding for ________?				
15.	Did ________ have visitors at home? a. friends other than group home members. b. family				
16.	a. Was there a behavioral program in effect for which data is on file? b. Did any new programs begin this week for which data is on file?				

17. Overall how would you rate your satisfaction with __________'s behavior this past week on a 5 point scale, with "5" being the most satisfied and "1" being the least satisfied? (By satisfaction, we mean to what extent has his/her behavior been problem-free and contributing toward group or family stability?).

18. Did anything unusual happen to or with ______________ this week?

Therefore, the TILE instrument provides data on the occurrence of one set of behaviors during each of 4 days during a 4-week period, the occurrence of another set of behaviors during each of the 4 weeks. Depending on the information needs of the data user, behaviors may be considered individually or combined into composites for more general descriptions of the subject's activities.

Method

Subjects and Residences

A total of 14 community residences for retarded adults in Oregon participated in the descriptive study. Of the 55 group residences in the state, 19 were selected for participation because they were located in towns with Oregon State System of Higher Education facilities. This selection eliminated long distance telephone charges for the study and resulted in a sampling of residences from throughout the state. Fifteen facilities out of the 19 agreed to participate in the study. In the four facilities that did not participate, two reported that not enough staff time was available to participate, one had no clients living in group residences, and one required too much time to gain informed consent for participation. One of the 15 facilities was dropped after the third week of the study when the respondent terminated employment with the agency. Two residents at each of the remaining 14 facilities were identified randomly from the population of each home. To insure random selection of subjects, the interviewer read a pair of numbers from a printed list of random numbers, and asked the facility director to identify, from a numbered list of residents' names, the two whose numbers matched those provided. For a description of the 28 subjects involved in the study see Table 2.

TABLE 2
Subject Characteristics

Descriptor	Mean	Range
Male Female Ratio (19/9)	32.6 yrs.	16 - 59 yrs.
Age		
Prior Institutionalization	12.2 yrs.	0 - 38 yrs.
No. of Residents	17.7	6 - 68
No. of Other Residents from Same Institution	10.9	0 - 22
No. of Total Staff	7.8	4 - 22
Time in Residence	30 mos.	3 mos. - 10 yrs.

Obtaining Participation and Consent

Prior to conducting the structured interviews, each community residence was contacted by telephone to request participation. Directors of the 14 facilities who agreed to participate were asked first to select two residents as described above, and then to select one staff member who spent the greatest amount of time with the two subjects. These staff members served as primary respondents during the four TILE calls. A second staff member was also selected to serve as a respondent during one TILE call for an independent agreement check.

Subsequent to this initial phone contact, Adaptive Behavior Scale materials and consent forms for both respondents and residents were sent to each facility. Interviews began for each facility after these materials were returned.

Interview Procedures

TILE data were recorded by three interviewers, two staff persons at the Specialized Training Program and one graduate research assistant. The same interviewer contacted the same facilities each week, at times selected by the respondents. Initial contacts and TILE interviews occurred between September 18, 1978, and November 3, 1978. Calls were between 10 and 20 minutes in duration. The interviewer identified himself/herself at the beginning of each call and immediately proceeded with the TILE interview. TILE questions were asked in the order they appear on the TILE form (see Table 1). Social conversation, lengthy questions, or discussions initiated by respondents during the interview were interrupted by interviewers and identified as appropriate at the conclusion of the interview. Questions regarding definitions of behaviors in TILE questions were answered by interviewers during the course of the TILE interview, using preprinted lists of behavioral definitions.

Obtaining Agreement Data

Agreement data were collected to check two potential sources of error with the instrument: that respondents' answers might not reflect actual events, and that records kept by interviewers might not reflect reported events. To assess respondent agreement TILE calls were placed to a second respondent at each facility,using the same procedures employed with primary respondents. Interviewer agreement was assessed with a second interviewer recording the respondents' answers during a TILE interview, while listening to the call on an extension phone. For each resident, agreement checks with second respondents and with second interviewers were scheduled for one of the four TILE calls. This was done by listing and numbering times of primary TILE calls to each interviewer and matching numbered call times to numbers obtained from a table of random numbers.

Results

The study provides both descriptive information on the group homes and residents who participated and methodological information on the telephone interview instrument. Descriptive data of interest include the percent of subjects who engaged in or experienced each of the target behaviors or program events, and more general descriptions using composite scores. Methodological data of interest relate to interviewer and interviewee agreement, reliability of scores, and concurrent validity. Finally, the instrument provides a basis for comparing lifestyle characteristics of residents of larger and smaller facilities in the sample.

Descriptive Data

Subject activities and experiences. Tables 3 and 4 provide descriptive data on 24-hour and weekly activities, respectively. The percent of subjects who did not engage in the behavior at all during the 4-week study is in the "no" column. The percent of subjects who did engage in the behavior during the 4-week study is found in the "yes" column. If subjects engaged in the behavior once during the month or more than once during the month, that is reflected in percentages found in those columns.

As can be seen in Table 3, the percentage of subjects who engaged in independent behaviors on a daily basis is very high: "Complete a chore independently" (100%) and "dress independently" (100%). The percentage of subjects who engaged in dependent behaviors on a daily basis "require help washing" (32%) and "require help toileting" (11%) is fairly low. The majority of subjects who required help toileting did so more than once during the month. This was not true of subjects who required help washing. The percentage of subjects who engaged in appropriate social behaviors is also high: "initiate social interaction" (89%) and "display affection" (82%). More than half of the subjects engaged in these social behaviors on a daily basis more than once during the month ("initiate social interaction" - 78%, "display affection" - 64%).

TABLE 3

Daily Descriptive Data: Monthly Summary

Items	no	yes	once during month	more than once during month
Argue	29%	71%	14%	57%
Complain	36%	64%	14%	50%
Initiate Social Interactions	11%	89%	11%	78%
Display Affection	18%	82%	18%	64%
Watch T.V. (more than 2 hours)	32%	68%	7%	61%
Display Stereotypical Behavior	43%	57%	7%	50%
Cry	82%	18%	4%	14%
Dress Independently	0%	100%	0%	100%
Require Help Washing	68%	32%	18%	14%
Require Help Toileting	89%	11%	0%	11%
Complete a Regular Chore Independently	0%	100%	0%	100%

A high percentage of subjects engaged in negative behaviors on a daily basis: "argue" (71%), "complain" (64%), and "display stereotypical behavior" (57%). More than three quarters of the subjects who engaged in these negative behaviors did so once during the month. The percentage of subjects who engaged in negative behaviors on a weekly basis was low: "steal: (18%), "leave home without permission" (18%), "fight" (32%), "destroy property deliberately" (14%), and "tantrum" (29%). None of the subjects destroyed property accidentally. Of the subjects who engaged in "fight" behavior,28% did so more than once a month. Of subjects engaging in other negative behaviors, 11% emitted "steal" behavior, 7% left "home without permission", 7% "destroyed property deliberately", 11% performed "tantrums" more than once a month.

Table 4 provides data on subject and staff behaviors that occurred during the weeks preceding each of the four TILE calls. The independent behaviors "fix something" and "use tools independently" were engaged in by a relatively small percentage of subjects (22%, 36%). More subjects performed social behaviors such as "community activities" (96%), "going out for meals" (79%), "shopping" (71%), and "field trips" (64%). "visits to friends" (50%) and by family and friends, "have visitors" (39%) were engaged in by somewhat fewer subjects. Twenty-eight percent of subjects who received visitors tended to receive them more than once during the month.

TABLE 4

Weekly Descriptive Data: Monthly Summary

Items	no	yes	once during month	more than once during month
Steal	82%	18%	7%	11%
Fix something	78%	22%	11%	11%
Destroy property accidentially	100%	0%	0%	0%
Use tools independently	64%	36%	4%	32%
Leave home without permission	82%	18%	11%	7%
Fight	68%	32%	4%	28%
Leave home for:				
community activity	4%	96%	7%	89%
family visit	58%	42%	21%	21%
visit to friend	50%	50%	21%	29%
meals	21%	79%	32%	47%
shopping	29%	71%	21%	50%
medical treatment	50%	50%	32%	18%
field trip	36%	64%	21%	43%
other	71%	29%	18%	11%
Destroy property deliberately	86%	14%	7%	7%
Tantrum	71%	29%	18%	11%
Staff turnover	25%	75%	25%	32%
Subject of staff meeting in house	43%	57%	25%	32%
Subject of staff meeting outside	78%	22%	11%	11%
Receive extra medical attendtion	54%	46%	25%	21%
Problem in funding	93%	7%	7%	0%
Have visitors	61%	39%	11%	28%
Program in effect w/data	32%	68%	0%	68%
New program w/data	58%	42%	21%	21%

Table 4 also indicates that 57% of the residents in this study were the subjects of in-house staff meetings, and 42% experienced implementation of new programs with evaluative data reported to be available. (Sixty-eight percent of the subjects had a program already in effect during the week preceding each call). Seventy-five percent of the subjects experienced a staff turn-over at least once during the month. Staff turn-over was defined as a change in staff due to termination and/or vacation.

Composites. These behavioral descriptions can be combined into composites for a more general description of an individual's behavior. Three composites, independent behavior, negative behavior, and social behavior, were generated to reflect aspects of individual lifestyle which might be of concern to program administrators, advocates, and evaluators.

The independent behavior composite consists of daily "dressing," "washing," and "toileting" independently, and "completing regular chores without a reminder." Weekly behaviors include "repairing some object," and "using tools independently." The negative behavior composite includes daily "arguing," "complaining," "engaging in stereotyped repetitive behaviors," and "crying." Weekly negative behaviors include "leaving home," "steal," "leave home without permission," "fight," "destroy property deliberately," and "tantrum." The social behavior composite includes daily "initiating social interactions," "appropriate affection," and weekly "activities away from home," and receiving visitors."

Scores for each composite were derived by summing the ratings for each behavioral category included, with non-occurrence counted as 0, occurrence once during the month counted as 1, and more than once scored as 2. Of a total possible sum of 48, scores for the independent behavior component ranged from 21-40 for all subjects. The mean score for independent behavior was 30.6, with a standard deviation of 6.6. Scores on the negative behavior composite, each of which could reach to a maximum of 72 points, ranged from 0-33 across subjects, with a mean of 9.7 and a standard deviation of 7.9. Of a maximum possible 88 points on the social behavior composite, subjects ranged from 4-53, with a mean of 24.3, and standard deviation of 12.9.

Methodological Information

In the previous section, data were presented for individual behavior occurrences and for behavior composites. For this reason, agreement percentages were computed for interviewers and respondents for both individual items and for composites. An agreement occurred when both interviewers or both respondents recorded or supplied exactly the same answer to a TILE question. All agreement percentages were derived by utilizing the following formula:

$$\frac{\text{Agreements}}{\text{Agreements \& Disagreements}} \times 100$$

Agreement between respondents. Interviewee agreement percentages for each TILE item are presented in Table 5. Interviewee agreement is high for each item except for "initiating social interactions" (61%) and "completing chores independently" (71%). Agreement is at 75% for "going out to meals," "going on field trips," and "watching T.V. for more than two hours". All of the other interviewee agreement percentages are 79% or higher, the majority of which (20) are 85% or higher. Interviewee agreement is also high for each TILE composite (75%-88%).

Agreement between interviewers. Interviewer agreement is consistently high for each TILE item ranging from 92%-100% (see Table 5). This is also true of interviewer agreement for TILE composites and for satisfaction ratings (range 98%-100%).

Reliability. Reliability in classical test theory concerns the consistency with which an instrument measures a trait or attribute. A reliable instrument is expected to result repeatedly in similar scores for given individuals relative to other individuals tested. Johnson and Bolstad (1973) suggest that evaluation of the reliability of behavioral measurement procedures is increasingly at issue as the variety of uses of behavioral data expands. The reliability of composite scores obtained was computed using the split-half method. Table 6 provides a summary of these split-half (Spearman-Brown cor-

TABLE 5

Inter-interviewee and Inter-interviewer Agreement Percentages

Daily Items	Respondent	Interviewer	Weekly Items	Respondent	Interviewer
Argue	82%	100%	Steal	93%	100%
Complain	79%	100%	Fix something	89%	100%
Initiate social interactions	93%	100%	Destroy property accidentally	100%	100%
Display affection	61%	96%	Use tools independently	79%	100%
Watch TV (more than 2 hours)	75%	96%	Leave home without permission	96%	100%
Display stereotypical behavior	82%	96%	Fight	69%	100%
Cry	93%	100%	Leave home for:		
Dress independently	96%	100%	community activity	89%	100%
Require help washing	96%	96%	family visit	89%	100%
Require help toileting	100%	100%	visit to friend	82%	100%
Complete a regular chore independently	71%	100%	meals	75%	96%
			shopping	79%	96%
			medical treatment	93%	100%
			field trip	75%	96%
			other	82%	100%
			Destroy property deliberately	96%	100%
			Tantrum	86%	100%
			Staff turnover	89%	92%
			Subject of staff meeting in house	89%	100% 100%
			Subject of staff meeting outside	86%	100%
			Receive extra medical attention	93%	96%
			Problem in funding	82%	100%
			Have visitors	89%	100%
			Program in effect with data	82%	92%
			New program with data	96%	100%

rected) reliability coefficients. Split-half correlations were performed by comparing the results of composite data collected during the first two and last two weeks. Split-half coefficients for the four TILE composites range from .71 to .92 (p < .01).

Concurrent validity. The concurrent validity of a test reflects the extent to which results agree with other measures of the same attributes or variables. High concurrent validity suggests that an instrument does measure what it purports to measure. The concurrent validity of TILE data was assessed by comparing composite scores with results of Adaptive Behavior Scale ratings and with general ratings of satisfaction. For this comparison, five scores were derived from the ABS results, based on factor analyses by Nihira (1969) and Lambert and Nicoll (1976). Given the composite descriptors, it was anticipated that significant correlation coefficients would be obtained between the independence composite and ABS1 (personal self-sufficiency), between the social behavior composite and ABS3 (personal-social responsibility), and between the negative behavior composite and ABS4 (interpersonal adjustment) and ABS5 (intrapersonal adjustment). As is apparent in Table 7, all the predicted correlations between TILE composites and ABS factors were significant at the .05 level. Further, the low correlations among the TILE composites and between TILE composites and unrelated ABS factors suggest that the independent behavior, social behavior, and negative behavior composites reflect different aspects of subject performance.

TABLE 6
Reliability Data for TILE Composites

TILE Composites/Rating

Reliability Descriptors	Negative Behavior		Independent Behavior		Social Behavior		Programming Decisions
Split-half Coefficient	** .71		** .84		** .92		** .88
	24 hr.	week	24 hr.	week	24 hr.	week	week
Percent Agreement Interviewer	96	99	98	98	100	98	99
Percent Agreement Interviewee	81	90	81	85	67	83	88

* significant at .05
** significant at .01

Additional support for instrument validity is provided by significant agreement between the general satisfaction rating (Item 17) and TILE composites, summed in all cases across all 4-weeks of data. The results of a stepwise multiple regression analysis indicate that 45 percent of the variance in the general satisfaction rating could be accounted for by scores in the three composites (R = .67). The negative behavior composite accounted for 32 percent of the variance (R = .56) while the social and independent behavior composites accounted for an

TABLE 7

Validity Coefficients for TILE Composites and ABS Factors and TILE Intercorrelation Coefficients

TILE Composites	ABS Factors					TILE Composites		
	Personal Self-Sufficiency	Community Self-Sufficiency	Personal-Social Responsibility	Interpersonal Adjustment	Intrapersonal Adjustment	Independent Behavior	Social Behavior	Negative Behavior
Independent Behavior	.44*	.31	.21	-.11	-.10		.28	.11
Social Behavior	.38*	.45*	.43*	.16	-.35			.09
Negative Behavior	-.39*	.24	-.43*	.49**	.76**			

* p = < .05
** p = < .01

additional 11 percent and 2 percent of the variance, respectively.

Large vs. Small Facilities

Table 8 presents a summary of ABS and TILE scores for the 28 residents, grouped by size of residence. The small facility group (less than the sample median of 15 residents) scored higher on every positive, normative measure for both ABS and TILE descriptors and lower on each negative, non-normative measure except ABS5 (intra-personal adjustment) where there is no difference between group means. The greatest differences between groups is found on ABS3 (personal-social responsibility) and ABS4 (inter-personal adjustment), where the large facility group (15 or more residents) demonstrated less personal and social responsibility and more social adjustment problems.

TABLE 8

ABS Factors and TILE Composites: Means and Standard Deviation for Large and Small Facilities

Size of Residence		ABS Factors					TILE Composites		
		Personal Self-Sufficiency	Community Self-Sufficiency	Personal-Social Responsibility	Interpersonal Adjustment	Intrapersonal Adjustment	Independence	Social Behavior	Negative Behavior
Small Facility (< 15)	x	78.8	56.5	48.1	21.6	3.6	32.9	27.6	7.7
(N = 14)	SD	4.8	14.2	8.5	21.8	3.4	5.3	13.8	5.8
Large Facility (> 15)	x	74.4	48.9	39.0	31.7	3.4	28.2	21.1	11.7
(N = 14)	SD	7.0	12.2	10.9	18.2	2.5	2.0	11.7	9.4

Discussion

This study provides a description of an assessment instrument designed to provide low cost data that are potentially useful to program evaluators and data users interested in life-style and quality-of-life data. An example of the use of TILE in the latter context is also provided with a description of the behavior and activities of 28 randomly selected retarded residents from group facilities in Oregon. The data reported are descriptive of both individual behaviors and differences between large and small group facilities on TILE and Adaptive Behavior Scale composites.

The reliability and validity data reported indicate generally high quality data for both single items and composites. Both interviewer and respondent agreements were consistently high. Split-half reliability coefficients are supportive of high internal consistency. Concurrent validity coefficients between selected TILE composites and ABS factors are .43 or higher and all are significant at the .01 level. Although additional validation data involving larger samples and extended observation periods should be obtained on the TILE instrument, the initial data presented in this study suggest that the three composites are valid behavioral descriptors for the 28 subjects sampled.

It is interesting to note that the satisfaction ratings supplied by residential care facility staff were more highly correlated with reported negative behaviors than with either social or independent behaviors. Apparently, the residential care facility staff involved in this study derived more satisfaction from subjects who did not engage in negative behaviors than from subjects who did engage in independent or social behaviors.

The daily item by item subject performances indicate that the majority of subjects engaged in independent and social behaviors. Further, subjects tended to engage in these behaviors during more than one 24-hour period preceding TILE calls. The weekly behavior percentages indicate some interesting staff behaviors. Staff turn-over was quite high during the four week study. This may be due to the inclusion of vacations in the staff turn-over definition. Relatively few staff meetings occurred between staff and staff persons outside the residential care facility. In addition, only slightly more than half of the subjects were subjects of in-house staff meetings during the four week study. An in-house staff meeting was defined as a staff discussion of an individual during a staff meeting within the group home with group home staff. Two-thirds of the subjects did have behavioral programs in effect with data, and close to half of the subjects experienced the initiation of new programs during the 4-week period. These data indicate that behavioral programs are a regular aspect of these facilities and that new programs are implemented frequently. Weekly data also indicate that almost 100% of the subjects engaged in community activities, and most of these subjects engaged in these community activities more than once during the month. The two other most frequent activities which subjects engaged in away from the facility were going out to meals and shopping. In addition, more than one-third of the subjects had visitors from outside the group home. Apparently, subjects in this sample entered into the community on a fairly regular basis.

The factor of residence size and the differences obtained in this study presents an interesting result in relation to previous studies. Bjannes and Butler (1974), in a study comparing two larger "board and care" facilities with two smaller "home care" facilities, found more independent behavior exhibited at the larger facilities. The present results indicate that the opposite is the case with the 28 subjects sampled: that is, small facility residents displayed more independent behavior than large facility residents, as shown in Table 8. Bjannes and Butler suggest that their results were a function of the differences in the amount of structure imposed upon the retarded residents. Data relating to this factor are not available in the present study.

Many potential applications of TILE exist. Quality-of-life data may be collected with TILE in conjunction with other instruments. TILE could be used to collect data to measure the effect of specific treatment interventions. Trends in TILE data on staff turn-over, medical treatment, and the implementation of behavioral programs may be of use to mental health professionals. Residential care facility staff may wish to collect and use TILE data to evaluate aspects of their programs, or to compare aspects of their programs with aspects of other residential programs. Finally, researchers may wish to further explore the effects on the behavior of retarded individuals living in small vs. large facilities, or institutionalized as opposed to community-based facilities.

References

Bell, N. J. IQ as a factor in community lifestyle of previously institutionalized retardates. Mental Retardation, 1976, 14 (3), 29-33.

Bellamy, G. T., Horner, R. H., & Inman, D. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Bjannes, A. T. & Butler, E. W. Environmental variation in community care facilties for mentally retarded persons. American Journal of Mental Deficiency, 1974, 78, 429-439.

Edgerton, R. Issues relating to the quality of life among mentally retarded persons. In M. J. Begab and S. A. Richardson (Eds.), The mentally retarded and society: A social science perspective. Baltimore: University Park Press, 1975.

Edgerton, R. B., & Langness, L. L. Observing mentally retarded persons in community settings: An anthropological perspective. In G. Sackett (Ed.), Observing behavior: Theory and applications in mental retardation (Vol. 1). Baltimore:

Gollay, E. Deinstitutionalized mentally retarded people: A closer look. Education and Training of the Mentally Retarded, 1977, 12(2), 137-144.

Irving, L., Crowell, F. A., & Bellamy, G. T. Multiple assessment evaluation of programs for severely retarded adults. Mental Retardation, 1979, 17(3), 123-128.

Johnson, S. M., & Bolstad, O. D. Methodological issues in naturalistic observation: Some problems and solutions for field research. In L. Hamerlynck, L. Handy, and E. Mash (Eds.), Behavior change: Methodology, concepts and practice. Research Press, Champaign, Illinois, 1973.

Jones, R. Observation by telephone: An economical behavior sampling technique. Oregon Research Institute Technical Report, 14(1), 1974.

Lambert, N. M., & Nicoll, R. C. Dimensions of adaptive behavior of retarded and non-retarded public school children. American Journal of Mental Deficiency, 1976, 2, 135-146.

Landesman-Dwyer, S., Stein, J., & Sackett, G. Group homes for the mentally retarded: An ecological and behavioral study. Report #2 of Developmental Disabilities Client Placement Study. Olympia: Department of Social and Health Services, 1976.

Landesman-Dwyer, S., Stein, J. G., & Sackett, G. A behavioral and ecological study of group homes. In G. Sackett (Ed.), Observing behavior: Theory and applications in mental retardation (Vol. 1). Baltimore: University Park Press, 1978.

Nihira, K. Factorial dimensions of adaptive behavior in adult retardates. American Journal of Mental Deficiency, 1969, 73, 868-878.

Connor, G. Home is a good place: A national perspective of community residential facilities for developmentally disabled persons. Monograph of the American Association of Mental Deficiency, 1976, 2.

NOLE:

The Naturalistic Observation of Severely Retarded Vocational Trainees in Non-Vocational Leisure Time Settings

Fred A. Crowell, Nancy Weissman-Frisch,
Deborah Lamson and J. Daniel Boomer

Preparation of this chapter was supported in part by a grant from the Bureau of Education for the Handicapped, Research Grant #G00-760-5455.

NOLE:

The Naturalistic Observation of Severely Retarded Vocational Trainees in Non-Vocational Leisure Time Settings

Deinstitutionalization, mainstreaming, and normalization represent current trends in the field of mental retardation (Nirje, 1969; Rosen, 1975). Devastating critiques of living conditions in institutions and a growing concern for human rights has shaped social consciousness to view institutionalization as inadequate and dehumanizing (Wolfensberger, 1972). Recent legislation has reflected this trend by prescribing guidelines for the delivery of vocational, (Vocational Rehabilitation Act, 1973) medical, (Wyatt v. Anderholt, 1971) and educational (P.L. 94-142) services. At the community level detailed standards for program development and evaluation are being established. Unfortunately, little empirical evidence exists to indicate exactly what services are needed to provide viable alternatives for the deinstitutionalized retarded individual. Thus, this massive move to deinstitutionalize has been described as a "...large-scale poorly controlled human experiment, guided by good intentions and economic considerations of politicians, planners, parents and professionals" (Landesman-Dwyer, 1976, p. 1). What is lacking is comprehensive assessment of the social/physical environment of the retarded. Data yielded through such assessment could provide an empirical basis for change as well as facilitate evaluating the effects proposed change could have on the lives of retarded individuals.

A major innovation in large-scale descriptive studies has been the use of naturalistic observation systems to obtain a more accurate and comprehensive picture of the complexity of the social environment of the retarded. As change agents, such a comprehensive understanding of the complex social system could facilitate our attempts at intervening and produce more effective interventions. Naturalistic observation may supplement other assessment procedures and provide important additional data for several reasons. First, naturalistic observation systems are generally high fidelity narrow bandwidth devices (Cronbach, 1970); that is, the focus of measurement is on high accuracy of detail across a small range of outcomes such as is required for pinpointing (the subtle) social contingencies. Second, use of naturalistic observation enables collection of data which is not dependent on the verbal skills of the individual being assessed; a skill deficit characteristic of the severely and profoundly retarded. The third advantage of naturalistic observation in program evaluation is the potential use in measuring both behavior and concurrent environmental events. Reliable monitoring of covariation over time between specific behaviors and stimuli can provide an empirical basis for many treatment decisions in program development and implementation.

Service delivery to severely and profoundly deinstitutionalized mentally retarded individuals requires that change agents carefully assess the existing social/physical environment. This will facilitate identification of individual needs, program development, and program evaluation. Naturalistic observation is one approach that yields objective, context-bound data while avoiding the potential bias of retrospective research techniques. Naturalistic observation, reliably used, can greatly increase our knowledge of the culture of retarded persons which, in turn, can form the basis for program development and evaluation.

The major purpose of this paper is to provide a description of a naturalistic observation instrument that was developed to assess the quality of life of severely retarded adults. The authors also provide descriptive data obtained during a continuing program evaluation effort involving vocational and community living skills training of severely retarded adults (Bellamy, Horner, and Inman, 1979).

NOLE: Naturalistic Observation of Living Environments

Instrument Description

The Naturalistic Observation of Living Environments (NOLE) system was designed to record the behavioral/environmental interactions of severely and profoundly retarded adults during unstructured, leisure-time activities. The NOLE system is a 10-second, fixed interval observation system consisting of 29 behavior categories and five setting categories. Observations are recorded every 10 seconds on a 2-page coding form (Figure 1) which contains: (1) a header section, (2) interval numbers (1-30), (3) space to identify the number of adults and peers present during each interval, (4) setting categories, and (5) behavior categories. The five setting categories and 29 behavior categories are defined in Appendix A.

Observer Recording Procedures

Observers record events every 10 seconds by observing subjects for 5 seconds and using the other 5 seconds to record observations. During each 5 second recording interval, hash marks are made through each symbol representing the coded behavior. Any changes in the number of adults or peers present or changes in settings are also recorded in the appropriate columns. Observers are synchronized to observe and record at 5 second intervals by means of a "beeper" tone emitted through earphones from the recording clipboard.

Subject______________ Date______ Observer______ Time______ Day________ Period________

Line	#A	#P	Setting	Att	Phy. Contact	Social Contact		Compliance
1			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1
2			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1
3			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1
4			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1
5			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1
6			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1
7			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1
8			D W R B W_1	S OB TV P A TV_1	Y OB PP S E	AG A P LA A_S P_S	DR SP	CO CO_1

See Appendix A for code categories.

FIGURE 1
NOLE Recording Form

Observer Training Procedures

NOLE observers are trained to reach an 80 percent agreement criterion for 15 minutes of observation on each of two consecutive days. The agreement percentage is computed for each category by dividing the number of agreements by the total agreements plus disagreements and multiplying the result by 100. A detailed description of the five phases of observer training is included in the NOLE Training Manual (Crowell, Lamson, Weissman-Frisch, & Bellamy, 1977).

NOLE Descriptive Data

Overview

Naturalistic observation codes have been used in a variety of settings but their use in providing descriptions of activities and environments of retarded persons has been quite limited (Irvin, Crowell, and Bellamy, in press). The descriptive data presented in the remaining sections of this paper specifically focuses upon social, independent and negative (non-normative) behaviors in a quality of life assessment context. In addition to the NOLE instrument, data will be presented from several rating instruments drawing upon the ratings of parents and caretakers.

Method

Subjects. The observation data reported in this descriptive study was collected on a total of 33 severely retarded adults at three training sites in the northwest which are replicating the vocational training model developed at the University of Oregon's Specialized Training Program. A description of subject characteristics in presented in Table 1.

TABLE 1

NOLE Descriptive Data: Replication Sites

Descriptor	Seattle	Reno	Spokane	All Sites
Age range	21-47	21-35	26-50	21-50
Mean age	31	26	35	31
I.Q. range	16-33	5-23	22-33	5-33
Mean I.Q.	24	14	31	23
No. of males	6	4	5	15
No. of females	5	5	7	17
Residence:				
Parent's home	6		1	7
Group home		9		9
ICF	3		11	14
Institution	2			2

Procedures. NOLE observations were made at four-month intervals on each subject for a 15-minute period on each of three consecutive days. The 45 minutes of observation were recorded during unstructured leisure time periods for each subject. During one 5-minute period for each subject, two observers recorded observations in order to obtain an estimate of observer agreement. At each 4-month observation period, data were also collected from significant others (parents, group home staff) who rated each subject on the Becker Adjective Checklist (BAC) and the Adaptive Behavior Scale (ABS).

Instrument Descriptors. For the purposes of data reduction and making statements about normative and non-normative differences or life style characteristics of these 33 retarded trainees, a number of composite descriptors were selected for each of the three instruments. A factor analytic study by Bellamy and Irvin (1977) on the BAC provided five factors used as descriptors in the present study. These factors were: withdrawn (B1), disruptive (B2), incompetent (B3), tense (B4) and non-compliant (B5).

A second five factors were selected for the ABS descriptors that combined a number of domain and sub-domain scores. The five ABS factors are: personal self-sufficiency (ABS 1), community self-sufficiency (ABS 2), personal-social responsibility (ABS 3), inter-personal adjustment (ABS 4) and intra-personal adjustment (ABS 5). The first three factors were identified in a study by Nihira (1969) and the last two factors were identified by Lambert and Nicoll (1976).

In addition to the single NOLE category of "aggression", four NOLE composites were selected as was a measure of "social complexity." The NOLE composites include: total deviant (TD), which combines categories AG, Y, and NC (non-compliance); total social interacts (TS), combining all adult (A), peer (P) and subject-initiated interacts (As, Ps); subject-initiated interacts (SII), which include both directive and supportive interactions; and environmental support, which includes the categories of adult support (ASP), peer support (PSP) and adult or peer compliance to subject-initiated directives. The social complexity (SC) descriptor consists of an information-theoretic measure of the variety of each subject's social behavior expressed as the observer's relative uncertainty (H_R) and is computed according to the following formula:

$$H_R = \frac{\Sigma - p \log_2 p}{H\max}$$

where "$-p \log_2 p$" is the average uncertainty measure for a given set of categories and "Hmax" is equal to $\log_2 N$, where N refers to the total number of social categories (Miller & Frick, 1949). The H_R measure, or SC as labeled in the present instance, can vary from 0 (total redundancy, stereotype) to 1 (maximum complexity, variety).

Results

The observer agreement percentages (computed by dividing the smaller number of category frequencies by the larger number of frequencies, summed across categories for each composite) are presented in Table 2 for the aggression category and the four composites. All five agreement percentages are within an acceptable range of 69 - 100% agreement, given the structural and functional complexity of the NOLE system. Table 2 also indicates high split-half reliability coefficients (Spearman - Brown corrected) for 17 workers at the Seattle and Reno sites. The coefficients were computed by comparing the first 22.5 minutes of an observation session with the second 22.5 minutes.

Table 3 presents descriptive data on the 33 subjects grouped according to type of residence, averaged across 1-3 (4-month) observation periods for each subject. Looking first at the NOLE categories, several differences between residence grouping are quite noticeable. Small facility (< 10 residents) subjects displayed more aggressive behavior than either of the other groups. Their total deviant rate (responses/minute) is also

TABLE 2

NOLE Categories: Observer Agreement Percentages and split-half Reliability Coefficients

NOLE Categories	Seattle	Reno	Spokane	Mean	split-half r
Aggression	98	92	100	97	.94
Total Deviant	73	65	70	69	.83
Total social	66	75	80	74	.84
Subject-initiated social contacts	76	80	76	77	.89
Environmental support	75	73	76	75	.85

higher than the family home and large facility groups (>10 residents). Although the social complexity index is somewhat higher than small facility residents, all three groups fall within a middle range of behavioral complexity. Comparative data on mild and moderately retarded persons on this measure would be interesting but is not currently available.

Table 3 also indicates that subjects who live in the large residential facilities initiate more social contacts than those who live with their families or in a small group facility. In terms of total social interactions and environmental support, however, the family home group and small facility group exhibits the highest rates.

TABLE 3

Descriptive Data on Three Types of Living/Leisure Environments

		NOLE						Becker Adjective Checklist					Adaptive Behavior Scale				
Residence Type		AG	TD	TS	SI	ES	SC	1	2	3	4	5	1	2	3	4	5
Family Home (N=7)	X	.001	.71	1.87	.56	.19	.452	-16.0	.9	-2.0	-1.6	-10.4	55.4	35.6	45.3	14.6	5.4
	SD	.004	[illegible]	.92	.77	.14	.2	6.0	9.1	5.4	12.8	5.7	8.3	7.8	6.2	12.3	5.0
Small Group Facility (N=9)	X	.07	1.26	.96	.50	.56	.588	-8.1	-2.6	.67	2.3	1.0	60.1	25.2	30.7	44.9	16.0
	SD	.15	.78	.65	.27	.27	.098	13.6	11.1	9.5	8.3	12.4	8.1	8.6	13.2	31.9	6.2
Large Group Facility (N=16)	X	.001	.46	.77	.41	.11	.469	17.1	-.13	2.9	-.11	1.0	63.0	26.1	34.3	23.7	4.4
	SD	.002	.66	.12	.15	.18	.245	8.1	8.9	8.1	7.2	8.8	19.5	11.8	11.6	19.5	3.7

As might be expected, the BAC and ABS mean ratings indicate that individuals living at home with families are seen as displaying more normative, problem-free behavior. The major differences between the large and small facility groups show up on ABS4 (inter-personal adjustment) and ABS5 (intra-personal adjustment). Small facility residents displayed more socially deviant and disruptive behavior than large facility residents, confirming the NOLE results for these same groups on the aggression and total deviant descriptors.

A second type of descriptive data, shown in Figure 2, compares the percentage of time that residents in the three types of living environments spent in total deviant, total social, and subject-initiated activities. The mean rates for each group of residents were converted into percentages by dividing each rate by the maximum rate of 6 responses per minute. The small facility residents were engaged in deviant activities almost a quarter of the time they were being observed compared with the family residence (12%) and the large residence (14%). Differences in total social and subject-initiated percentages were minimal.

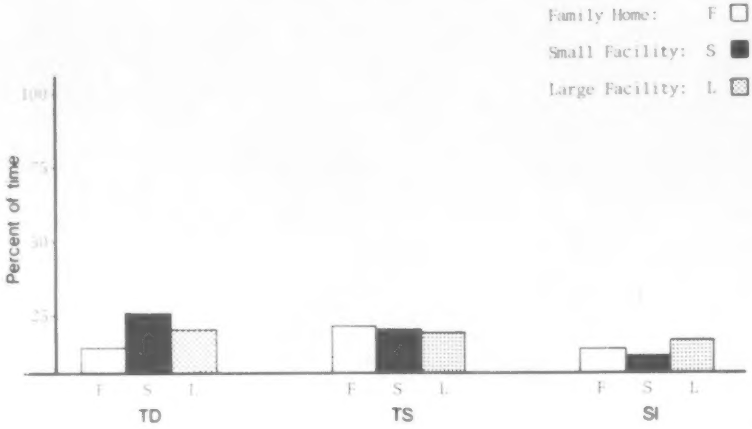


FIGURE 2

Percent of Time in Total Deviant (TD), Total Social (TS) and Subject-initiated (SI) Activities for Three Living Environments

Discussion

The data generated by the NOLE system in this report is descriptive of severely retarded adults within three types of residential facilities. The data presented revealed behavioral differences between these settings on several composite NOLE variables that were essentially replicated by similar measures on the Becker Adjective Checklist and the Adaptive Behavior Scale.

The present study confirms the results of a study by Bjaanes and Butler (1974) comparing two larger "board and care" facilities on the variable of "independence." In both cases the large facility residents demonstrated more independent behavior. Additionally, the present results indicate that the severely retarded adults observed in a family environment are more independent than those observed in the two types of group facilities. While suggestive of setting differences, the present data is based upon a small number of residents at only a few residence facilities (only one in the case of small group facility). The data are presented as a means of providing an example of one application of the use of the NOLE instrument in asking questions about quality of care or quality of life issues in designing service delivery programs for retarded persons.

Quality of Life Assessment

With increasing emphasis upon deinstitutionalization, normalization, and community-based programs, there is a growing concern about the assessment of quality of life of retarded persons (Edgerton, 1975: O'Connor, 1976; Landesman-Dwyer, Stein, and Sackett, 1976). Evaluation instruments are needed which provide program managers with descriptive data on life styles and activities of retarded persons in a variety of environmental settings. The NOLE observation system described in this report is a potentially useful instrument for providing quality of life data on activities and behavior of retarded persons in living/leisure environments.

The rationale for the selection of the NOLE categories and composites is based upon a conceptualization of "quality of life" that stresses increased participation in vocational and social activities. Such participation is incompatible with stereotypical, negative behaviors including aggression and non-compliance. Increased participation is also assumed to increase the complexity of the behavioral repertoire as indexed by the NOLE categories. Finally, it is theorized that QOL is enhanced by increased support from an individual's environment. Given the NOLE categories one can develop a number of composite measures, such as those used in the present study, to assess QOL of retarded persons in a variety of settings.

Given the rationale for QOL assessment, how does the present data inform us about the life-styles and quality of living/leisure environments of retarded persons? The image generated by the data is one of an environment that is not very supportive or responsive to severely retarded persons. It is an environment where participation in social activities and decision-making is extremely limited as indexed by the number and kinds of interactions observed across residences and subjects. Positive, supportive physical contacts occur infrequently when initiated by either the subject to adults and peers or when the latter initiate such contacts to the subjects. Although aggressive behaviors are low probability behaviors, stereotypical behaviors are frequent enough to be incompatible with the subject's participation in many social activities.

Although naturalistic observation instruments such as NOLE can and should be used in program evaluation applications, the data that is generated is necessarily of a single perspective type. Quality of life assessment, in either a descriptive sense or as a measure of program impact, demands multiple measures, combining complementary perspectives (Crowell, in press). When combined with measures representing other perspectives the NOLE system may be useful in contributing to our total pool of knowledge about the quality of the retarded person's environment, the complexity of which is too often reduced to match the simplicity of one's single perspective assumptions. Edgerton (1975) has stated the case most persuasively:

> But then science itself, at least according to some, is not the reduction of complexity to simplicity, but the process of understanding complexity in its own right.
>
> I say all of this because I have come to realize that there is far more to mentally retarded persons than I had previously realized. My own writing about mentally retarded persons does not seem to me wrong, it is simply unfinished. Not only have I not got to the bottom of things, I cannot even guess where the bottom is. So I end with a cautionary note, for myself, if no one else: mentally retarded persons, even though they may have lesser skills and capacities than the rest of us, are nevertheless complex persons who live immensely complicated lives, like the rest of us (p. 137-138).

References

Bellamy, G. T., Horner, R. H., & Inman, D. P. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Bellamy, G. T., & Irvin, L. Utility of an instrument for assessing labels to describe retarded adults. In G. Bellamy, R. Horner, and D. Inman (Eds.), Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program, Vol. 2. Eugene, Oregon: University of Oregon Center on Human Development, 1977.

Bjannes, A. T., & Butler, E. W. Environmental variation in community care facilities for mentally retarded persons. American Journal of Mental Deficiency, 1974, 78, 429-439.

Cronbach, L. J. Essentials of psychological testing. New York: Harper and Row, 1970.

Crowell, F. A. Evaluating programs for educating mentally retarded persons: Changing paradigms. In J. Gottlieb (Ed), Educating mentally retarded persons. Baltimore: University Park Press, in press.

Crowell, F., Lamson, D., Weissman-Frisch, N., & Bellamy G. *NOLE: A naturalistic observation study of living environments of severely retarded vocational trainees.* Unpublished manuscript, University of Oregon Center on Human Development, Eugene, 1977.

Edgerton, R. Issues relating to the quality of life among mentally retarded persons. In M. Begab and S. Richardson (Eds.), *The mentally retarded and society: A social science perspective.* Baltimore: University Park Press, 1975.

Irvin, L., Crowell, F. A., & Bellamy, G. T. Multiple assessment evaluation of programs for severely retarded adults. *Mental Retardation*, in press.

Lambert, N. M., & Nicoll, R. C. Dimensions of adaptive behavior of retarded and non-retarded public-school children. *American Journal of Mental Deficiency*, 1976, *2*, 135-146.

Landesman-Dwyer, S., Stein, J., & Sackett, G. *Group homes for the mentally retarded: An ecological and behavioral study.* Report #2 of Developmental Disabilities Client Placement Study. Olympia: Department of Social and Health Services, 1976.

Miller, G. A., & Frick, F. C. Statistical behavioristics and sequences of responses. *Psychological Review*, 1949, *56*, 311-324.

Nihira, K. Factorial dimensions of adaptive behavior in adult retardates. *American Journal of Mental Deficiency*, 1969, *73*, 868-878.

Nirje, B. The normalization principle and its human management implications. In R. Kugel and W. Wolfensberger (Eds.), *Changing patterns of residential care for the mentally retarded.* Washington, D. C., PCMR, 1969.

O'Connor, G. *Home is a good place: A national perspective of community residential facilities for developmentally disabled persons.* Monograph of the American Association of Mental Deficiency, 1976.

Rosen, D. *Deinstitutionalization.* Invited presentation at the Conference on the Severely Handicapped, University of Oregon, February, 1975.

Wolfensberger, W. *The principle of normalization in human services.* Toronto: National Institute on Mental Retardation, 1972.

Wyatt vs. Aderholt, 344 F. Supp. 1341 (N. D. Alabama, 1971) 324 F. Supp. 781 (N. D. Alabama, 1971).

Appendix A

BEHAVIOR CODE CATEGORIES

Columns II. and III: The observer will maintain an ongoing record of the number of nonresident adults (#A) and the number of resident adult peers (#P) by counting and recording this number in their respective columns. At the onset of the observation period the observer will recode with each addition or subtraction of peer or adult.

Column IV: The location of the subject being observed is described in terms of 5 optional behavior settings. Select that option which best describes the setting at the onset of coding. Code again when either the setting changes or when beginning a new coding sheet.

- D: Dining is coded when the subject is engaging in activities related to eating and/or is within a setting designated for dining.
- W: Work is coded when the subject is engaging in discrete type tasks. Discrete here refers to behaviors that involve making a single change in the present environment, e.g., dusting (from dusty to not dusty) and sweeping and/or within a setting designated for work related activites.
- W_1: $Work_1$ is coded when the subject is engaging in continuous or combination of continuous/discrete tasks such as dusting then sweeping then doing the dishes; a series of related tasks; within a setting designated for work related activities.
- R: Recreation is coded when the subject is in a setting designated for leisure type activities. This category will include informal living room type settings and/or formalized recreational areas such as game rooms, TV rooms, etc.
- B: Break is coded when the setting is characterized by transition from one activity (such as eating) to another (such as watching TV).

Column V: Attention. The apparent focus of the subject's visual attention is coded each 10 second interval.

It may be focused on:

- S: Self denotes attending to subject's own body (looking at one's hand or adjusting clothing).
- Ob: Object denotes attending to any (and all) inanimate objects (wall, toy, furniture) and any (and all) non-human animate objects (animals, plants,

etc.) within the environment.

TV: Television denotes attending to the television discretely; looking at TV-look away, look back (duration-3 seconds or less).

TV_1: Television denotes attending to the television with dual modality, looking and listening with a continuous response (duration - 3 seconds or more).

Column VI: The non-verbal physical interaction between the subject and the environment may be coded:

Y: Stereotypical behavior which denotes that the subject is engaging in repetitious body movements or fixating on a single body part, which are judged by the observer as non normative. Y may be coded simultaneously with other behavior codes.

S: Self denotes normative touching of body parts.

OB: Object denotes subject object interaction characterized by object function and/or normative non-repetitious continuous motions.

PP: Positive Physical denotes a physical show of affection toward another person in what is judged as in a normative and positive way.

E: Denotes a physical movement involving the subject's total body, from one spatial position to another.

Column VII: The verbal, vocal or gestural interaction between the subject and another person may be coded:

AG: Aggression which denotes physically aggressive acts directed at self, objects and other individuals

A: Adult denotes an adult initiated social interaction which includes non-directive, directive (double coded with DR) and supportive (double coded with SP) statements.

A_S: $Adult_S$ denotes a subject initiated social interaction directed toward a non-resident adult.

P: Peer denotes a peer initiated social interaction which includes non-directive directive (double coded with DR) and supportive (double coded with SP) statements directed toward the subject.

P_S: $Peer_S$ denotes a social interaction initiated by the subject and directed toward a peer.

LA: Laugh denotes vocal laughs or sustained smiling by the subject. LA can be double coded.

Column VIII: All adult or peer initiations are coded as they occur with only one initiation coded per 10-second interval. These will be coded as:

SP: Supportive which denotes verbal or physical interactions directed at the subject by either an adult or a peer, e.g., verbal praise, encouragement, conversation, smiling. Supportive statements imply that the adult or peer approves of the subject's behavior and expects no changes in that behavior.

DR: Directive denotes verbal or physical interactions directed at the subject by either an adult or peer in which the subject is directed to change activities, is corrected for her/his behavior, is criticized, or is asked a question.

Column IX: Following all directives from adults or peers the subject's behavior will be coded CO (compliance). Non-compliance is not coded but calculated by subtracting the number of CO's from the number of directives given. The coder will do this at the completion of the observation session.

CO: Denotes compliance to a directive from an adult or peer.

CO_1: Denotes compliance of the adult or peer to the question, request or other directive initiated by the subject.

NC: Denotes non-compliance to an adult or peer-initiated directive.

NC_1: Denotes adult or peer non-compliance to subject's directive.

Procedural Category: In all coding intervals where the view of subject is obstructed (moving to another room, behind a large object, etc.) the observer places a check mark (✓) in the attention column.

The Trainee Performance Sample: Improving Vocational Skill Assessment of Severely Retarded Adults

Larry K. Irvin, Russell Gersten, G. Thomas Bellamy,
Valerie Taylor and Daniel W. Close

The Trainee Performance Sample: Improving Vocational Skill Assessment of Severely Retarded Adults

Despite the general tendency within special education to question the instructional utility of norm-referenced tests (e.g., Popham, 1974; Salvia & Ysseldyke, 1978), increasing deinstitutionalization of severely/profoundly retarded adults has created exactly that situation where norm-referenced tests can serve appropriately: the selection and placement of individuals in a reasonably objective, empirically validated fashion (Anastasi, 1976). Yet, a recent survey of vocational assessment instruments currently available for this population yields rather disheartening results (Halpern, 1978). Despite the great strides made in providing successful vocational training for severely retarded individuals (e.g., Bellamy, Horner & Inman, 1979; Gold, 1973), little has been achieved in the way of vocational assessment instrumentation for this population. The general tendency of Bellamy et al. (1979) and others (Gold, 1974; Haywood, 1975) is toward idiographic task-specific behavioral assessment. While this methodology certainly is superior to traditional norm-referenced testing for ongoing analysis of individual worker performance, it does not ameliorate the problems raised by deinstitutionalization--namely, to place large groups of severely retarded adults in appropriate training programs.

The type of instrument typically advocated for vocational assessment is the work sample (Kulman et al., 1975), "a well defined work activity involving tasks, materials and tools which are identical or similar to those in an actual job or cluster of jobs" (Kulman et al., p. 55). Despite its surface appeal, there are numerous problems in applying work sample methodology with severely retarded persons. For one thing, as Halpern (1978) observes, work samples are often much more highly contrived than they purport to be; few capture the essence of on-the-job work performance. In addition, the validity of work sample assessment has not been established. Rosenberg's (1967) predictive validity studies of the TOWER (Testing Orientation and Work Evaluation in Rehabilitation, 1974), one of the most highly developed systems of work samples, demonstrated extremely weak, nonsignificant relationships between TOWER scores and worker performance.

Even if there were empirically validated work samples available for retarded persons, they would be of dubious value for severely retarded adults for several reasons. First, work samples measure the products of prior learning. Individuals with impoverished educational and training backgrounds are heavily penalized on product measures (Irvin & Halpern, 1979). Recently, a process approach has been proposed as an assessment methodology with severely retarded persons (Budoff & Hamilton, 1976; Feuerstein, 1972; Irvin & Halpern, 1979). In process assessment, an individual's ability to benefit from instruction (or training) is measured. Initial results with EMR students indicate greater predictive validity for process measures than conventional IQ tests (Budoff & Hamilton, 1976). Also, Cromwell's (1976) review of predictive validity in the psychiatric field indicated that assessment of current skill level is often not a potent predictor of future performance, unless supplemented by other information on the client and relevant environmental settings.

Another problem with work samples is that they adopt a macroscopic approach toward evaluating performance. Current practice demonstrates that successful vocational training for severely retarded workers depends on the microscopic breakdown of large tasks into small components so that chains of behavior can be taught (Gold, 1973; Bellamy et al., 1979).

A final concern regarding work samples is that the criterion behavior they purport to predict, ultimate level of productivity, may not be the most appropriate criterion variable, especially in light of accumulating evidence of the effects of task and situation variables on both learning and productivity of severely retarded persons. It seems more sensible to focus on the amount of time required for training an individual, given a detailed training procedure and a specific work task. This approach focuses attention on the fact that an individual's level of vocational achievement is not static, but can be altered through appropriate training.

With this approach the question for evaluation becomes one of predicting training time needed to learn a specific task. The first step is, of course, to identify individual attributes which correlate with this criterion. A related assessment question is whether individual attributes interact with specific variations in training procedure, so that different procedures are more efficient with different individuals. The major purpose of our research derives directly from these concerns and involves development of a vocational assessment instrument for use with severely retarded persons that:

1. Uses brief tasks as test items, rather than longer work samples;
2. Uses a process rather than a product approach, i.e. provides measures of an individual's ability to benefit from current training rather than prior knowledge; and
3. Utilizes training time for specific tasks, rather than ultimate client productivity as criterion measures.

One approach to vocational assessment instrumentation for use with severely retarded persons which incorporated some of these features was conceived by Bellamy and Snyder (1976). Their Trainee Performance Sample was designed to predict training time (and cost) for severely and profoundly retarded adults. The underlying assumption was that severely retarded adults can be trained to perform complex vocational tasks such as construction of cable-harness assemblies (Bellamy et al., 1979), or bicycle brakes (Gold, 1973), but that the amount of training time necessary would be variable from individual to individual. The underlying premise was that if a test was constructed of items that sampled in a representative fashion the domain of tasks commonly used in vocational training for severely retarded individuals, the composite score should have some predictive validity. Items were selected that paralleled many of the components of the Specialized Training Program, a vocational training program for severely retarded adults. Items were selected which involved tasks requiring visual discrimination of minimal differences, two-hand coordination, and sequencing of task components (see Figure 1 for examples).

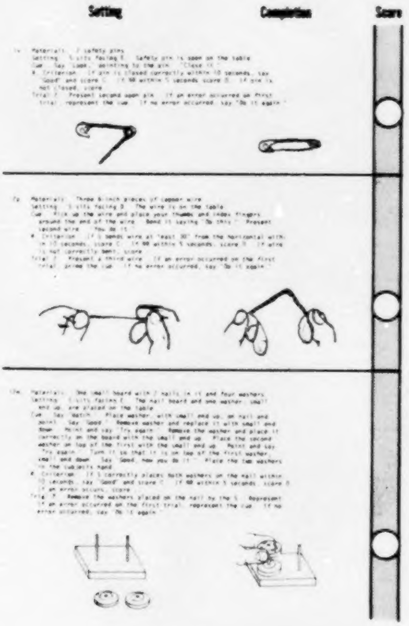


FIGURE 1

Sample Items from Trainee Performance Sample

The two unique features of the initial version of the TPS were: Strict adherence to the use of specific versus general norms (Parker & Hansen, 1976); and modified incorporation of the learning potential model (Irvin & Halpern, 1979; Feuerstein, 1972). To meet the first criterion, the test was normed exclusively on a sample of severely/profoundly retarded adults at a state institution in the Northwest. Incorporation of a modified learning potential model was the most innovative feature of the test. Each of the 30 items included a training procedure within the test item. Two types of procedures were used--modeling/match-to-sample techniques (on 26 items) and physical prompts (on the remaining 4 items). Thus, to some extent, the initial version of the TPS measured ability to benefit from instruction, rather than prior knowledge.

The TPS also incorporated a correction procedure within the format. If an individual correctly performed the item on the first trial, s/he was simply asked to "do it again". If s/he was unsuccessful the first time, the test administrator repeated the model (or prime) offering the examinee an opportunity to benefit from a correction procedure (see Figure 1). Individuals who performed correctly after correction received 1 point, those who met the criterion on both trials received 2 points, and those who were unsuccessful both times receive 0 points. Total scores on the TPS as initially standardized ranged from 0 to 60 ($\overline{X}$ = 30.8, sd = 17.6). Psychometric analyses demonstrated a remarkably high internal consistency (Coefficient α = .961), and high test-retest reliability (r_{tt} = .91). Exploratory studies in criterion-related validity appeared promising. There were reasonably strong appropriately negative correlations between the TPS scores of 111 institutionalized adults and the training time necessary to teach several vocational assembly tasks (such as a 3-piece heat-sink assembly) via a variety of types of chaining procedures (median r = -.72). However the relatively small sample size per task, and the very limited domain of small assembly tasks selected, limits the generalizability of the validity findings.

In 1977, a decision was made to revise the Trainee Performance Sample. The major purpose of the revision was:

1. To increase the range of potential appropriate examinees to include individuals classified as low-moderate mentally retarded;
2. To sample, in a systematic fashion, type of task and instruction dimensions; and
3. To conform to APA test development guidelines regarding reliability and validity.

The original TPS (Bellamy & Snyder, 1976) emphasized the use of instruction by models, with only a few items involving physical prompts and none involving verbal directions only. The articulation of a comprehensive vocational training methodology for severely retarded individuals (Bellamy, Horner & Inman, 1979) enabled a more systematic approach to the generation of test items.

Method

Instrumentation

Content validity. Content validity has become an increasingly important concern in psychometric thinking in the past decade (Anastasi, 1976; Becker & Engelmann, 1976; Salvia &

Ysseldyke, 1978). This reflects two current trends in field of measurement. The first is a concern with linguistic validity (Payne, 1975), an attempt to clarify exactly what is meant by terms like "reading comprehension ability," "vocational competence," and "mathematical aptitude." The second is the tendency towards the use of criterion-referenced tests (Popham, 1974), or domain-referenced tests (Hively, 1975), where the exact domain of objectives or behaviors to be measured is clearly specified. Halpern (1978), among others, has indicated some of the drawbacks of using criterion-referenced tests in vocational assessment, i.e., the difficulties in specifying everything of importance as a behavioral objective, and the lack of a method for assuring that the specified domain has real predictive validity. These arguments seem particularly appropriate for a relatively global measure like the Trainee Performance Sample. The revision of the TPS involved efforts to incorporate some of the innovative ideas of the developers of criterion-referenced tests within the framework of conventional norm-referenced assessment. That is, item types are more analogous to typical criterion-referenced test items, but interpretation of individual performance corresponds more closely to traditional norm-referenced approaches. A major difference between the two versions of the TPS is a clearer methodology for item selection on the revised TPS, which incorporated concepts from both domain-referenced tests (Hively, 1975) and instructionally-based tests (Becker & Engelmann, 1976). An item generating grid was developed based on the vocational training methodology of the Specialized Training Program (Bellamy, Horner & Inman, 1979). Figure 2 illustrates the three components of the grid-type of instruction, task attributes, and number of objects involved in the task. The grid was used as a basis for generation of items. Sometimes the demands of feasibility and logic made some cells impossible to fill. The three types of instructions--model/match-to-sample, physical prompts, and verbal directions are the three types of training procedures employed in vocational training for severely retarded adults.

Use of the grid insured that one-third of the test items would employ each procedure. The task attributes--minimal difference, coordinated two-hand movement, if-then rule, sequencing, conjunctive stimulus control--involved out of Bellamy, Horner, and Inman's (1979) task analyses of vocational tasks. Bellamy et al., believe that these might serve as difficulty descriptors for vocational tasks. Use of the grid also insured that each of these task attributes was representatively sampled across both instructional formats and tasks with varying number of component objects. In addition, one-sixth of the items were selected to have none of these attributes present, presumably "easy" items. The third dimension of interest in item development/revision was number of objects within each task. This component has less of a theoretical rationale; rather, it serves to guarantee a variety in types of test items. Finally, to insure that the verbal instructions were not beyond the receptive language capacities of the population, a Basic TPS Vocabulary List (Figure 3) was designed. The grid and the vocabulary list were adhered to as much as was logically possible. Certain grid entries were either logically impossible, or instructionally

	Single	Double	Movement
Verbal	3 items	6 items	1 item
Model	3 items	6 items	1 item
Prime	3 items	6 items	1 item

Dimension 1

1. Verbal Directions
2. Model/Match-to-Sample
3. Physical Prime

Dimension 2

1. Manipulation on Single Object = SINGLE
2. Manipulation on 2 or More Objects = DOUBLE
3. Physical Movement - or Position: No Objects = MOVEMENT

Difficulty Descriptors -- Dichotomous -- (Sampled representatively across instructional format and number of objects dimensions)

0 = None Present
1 = Minimal Difference
2 = Coordinated (2-hand) Movement
3 = If-Then Rule
4 = Conjunctive Stimulus Control (i.e., 2 Stimulus - Components)
5 = Sequence of Movements Required

FIGURE 2

Item Blueprint for Generating TPS Items

unsound (such as teaching visual discrimination by the use of physical primes). A total of 31 revised or newly developed items were piloted.

Concurrent validity. The search for instruments to ascertain concurrent validity was problematic because of the lack of vocational rating scales appropriate for a severely handicapped population. The two most likely contenders, the San Francisco Vocational Competency Scale (Levine & Elzey, 1968) and Palmer's (1974) vocational rating scale, were deemed inappropriate. Though both possessed adequate reliability and validity characteristics (Buros, 1972), they were clearly geared for use with mildly handicapped persons; many items were not relevant to activities engaged in by severely retarded workers.

Palmer's (1974) factor analysis of his checklist was used as a basis for insuring broad, representative sampling of behaviors in a newly created supervisor's checklist which included only behaviors appropriate for severely/profoundly retarded individuals. Palmer found four major categories of work behaviors which correlated with vocational success: stability of worker's performance ability to follow instructions, social/cooperation and peer relation skills, and coordination and physical dexterity skills. Palmer's categories and sub-categories were used as a guide for generating items from the vocational training staff of the Specialized Training Program. These items were then collated, rewritten, and organized into a concurrent validation scale in a manner which insured that

Palmer's (1974) categories were adequately sampled. Each item was a behavioral statement such as, "Responds to simple verbal directions without requiring further assistance" or "performs tasks which require the use of a hand tool, such as a screwdriver, with minimal assistance." Each statement was followed by a 3-point rating scale, 19 of which involved frequency ratings ("usually or always," "sometimes," "seldom or never") two of which involved independence of skill ratings ("independently," "with some assistance," "not at all") and two of which involved speed ratings ("rapidly," "at a moderate rate," and "slowly or not at all").

Nouns	Verbs	Adjectives and Adverbs	Prepositional Phrases
It, these, this	Are	Size - smallest	In
block	close	color - red	on top of
colors	covers	adjective - both, same	under
hand	put	adverbs - over, when	up together
hole	screw		then
lid	stand		
paper	stop		
penny	touches		
side	turn		
table			

FIGURE 3

TPS Basic Vocabulary

Sample and Procedure

Delineating an operational definition for the sample proved to be problematic. The most appropriate population for tryout of the revised TPS was adults classified as severely retarded, profoundly retarded, or those in the low moderate range, i.e., those individuals with an IQ of 45 or below. However, many of the adults in institutions and work activity centers had not taken IQ tests in years. Furthermore, the appropriateness and accuracy of a WAIS or Stanford-Binet score for a severely/profoundly retarded person is questionable. Rankin (1978) has indicated that the major purpose of these tests is to screen out individuals with low IQs. Because people already in state institutions or work activity centers are already effectively screened, administering these tests would have served no purpose. In addition, the difficulty level and high verbal loading of standardized adult intelligence tests make administration a frustrating and not very meaningful experience for severely retarded persons (Anastasi, 1976). An examination of IQ tests with a lesser emphasis on expressive language--such as the Peabody Picture Vocabulary Test, and the Leiter Scale--revealed that these tests were normed only on children 16 or younger. Thus, a decision was made to define the population in the following manner:

1. In general, a "community-based definition" of profound-severe-low moderate retardation was used (Mercer, 1974). Individuals so classified by the staff of work activity centers would be considered eligible candidates for the TPS. If a potential examinee received full credit on the basis of per-

formance on a pre-test which consisted of five items like those on TPS, the individual was regarded as too high functioning for the TPS to produce a useful score. TPS testing was not attempted for such individuals. Generally, the activity center staff classifications were quite accurate. Only a few were not tested on the TPS because of pre-test results;

2. Raw scores on the Peabody Picture Vocabulary Test (Dunn, 1973) were recorded for all individuals. They provided easily replicable descriptions of the sample. Throne (1972) argues that raw scores can offer a reasonably reliable description of the range of verbal-intellectual abilities of a test sample. The fact that the Peabody Picture Vocabulary Test (Dunn, 1973) has fewer items than other IQ tests beyond the cognitive grasp of severely retarded adults, makes it a reasonable instrument to use for sample description purposes.

Thus, 149 adults classified as profoundly retarded, severely retarded or moderately retarded were selected as participants in May, 1978. Fifty-four percent were enrolled in community work activity centers and 46% in residential institutions. All were over the age of 16, with no major physical impairments or visual impairments. A sample of 26 was retested two weeks after the original test. All participation was voluntary, with referrals originating from cooperating residential/work facilities, and final informed consents coming directly from participants and/or their legal guardian(s).

All participants were tested on the revised TPS by project staff during May through July, 1978. Testing required one 30-minute session per participant, and occurred at the participant's residence/work site.

To determine the extent to which APA test development criteria were met, several procedures were employed. After the Trainee Performance Sample was administered, Coefficient Alpha (Nunnally, 1967) was calculated on both the full TPS score and two sub-scales (verbal instructions, nonverbal instructions). This index is generally believed to provide the best measure of internal consistency for tests containing items of variable difficulties. Temporal stability was estimated by calculation of a conventional test-retest correlation (Anastasi, 1976) on the 26 individuals who were administered the TPS twice.

To determine concurrent validity, examinees' revised TPS scores were then correlated with full scores on the newly created Supervisor Checklist. The most appropriate work supervisor completed the Supervisor Checklist for each TPS examinee. Directions included an instruction to rate individuals' performance within the context of the performance of other individuals at that work site. Completion required about 5-10 minutes per examinee.

An additional measure of concurrent validity, albeit a tentative one, is correlation of TPS scores with PPVT raw scores. This provides a rough estimate of how TPS scores correlate with scores taken from a standardized intelligence test. The fact that the Peabody Picture Vocabulary Test was not used on the population on which it was normed (children under 16) makes this a tentative estimate. PPVT testing was accomplished by project staff in a 5-10 minute session with individual examinees just prior to TPS testing.

Results

The major research efforts have been in the areas of: (1) test construction and content validity; (2) internal consistency, temporal reliability and item analysis; and (3) concurrent behavioral validity.

Content Validity

The content validation results have been described within the "Methodology" section of this report, and will not be further detailed here. Since only one item has been eliminated after item analyses, all levels of all three dimensions in the TPS item generation blueprint are represented in the final test (see Figure 2).

Revised TPS Difficulty, Variability, and Reliability

Mean difficulty level of the 30-item revised TPS was 35.5/60 or 59% correct, for the total sample of 149 examinees. The standard deviation was 15.6. Average difficulties of the "subscales" were: verbal only--56%, model and/or prompts--61%. Standard deviations of the subscales were 5.4 and 5.6, respectively.

Item to total test correlations were computed for each item. These are analogous to the traditional point-biserial correlations calculated for achievement tests (Anastasi, 1976), except that the range per item on the TPS is 0 through 2, rather than right vs. wrong. The Pearson r's ranged from .20 to .78, with 21 items having correlations in the .50 to .69 range.

Coefficient Alpha internal consistency reliability for the 30 item revised TPS was $r_\alpha = .95$. Test-retest reliability for total TPS scores was $r = .93$ for $N = 26$. Means and standard deviations at initial and later (two-three weeks) testings were equivalent: $\bar{X}_1 = 41.4$, $\bar{X}_2 = 41.2$; $sd_1 = 12.7$, $sd_2 = 12.6$.

The 10-item verbal subscale (items with verbal directions only) also demonstrated considerable internal consistency reliability with $r_\alpha = .84$. The 20-item model/prompt subscale had an even higher reliability coefficient, $r_\alpha = .94$. The correlation between item score and subscale score was computed for each item; the range was from .25 to .78 for the verbal items, and .38 to .78 for the model/prompt items.

Overall, as is evident, the reliabilities, difficulty levels, and variabilities of the revised TPS and subscales are appropriate. In each of these respects, the TPS exceeds APA (1974) standards for test development.

Concurrent Validity

Concurrent validity of the TPS was assessed in three different ways:

1. Correlation of TPS scores with supervisor checklist ratings;
2. Correlation of TPS scores with raw scores on the Peabody Picture Vocabulary Test; and
3. Contrasting the distribution of TPS scores for institutionalized participants with the distribution of scores for community-based participants (Anastasi, 1976).

Prior to validity analyses, the psychometric characteristics of the Supervisor Checklist were evaluated. Mean ratings, variability of ratings, and Coefficient Alpha internal consistency reliability of ratings were calculated. Mean rating of performance on the total checklist was 50.9 for 23 items. The resulting average item rating was 2.21 out of a possible 3, slightly above the midpoint value. Since supervisors were asked to rate each worker compared to other workers at that shop, the expected mean per shop would be 2.0. The deviation between actual mean and expected mean is only .2, indicating a very slight tendency towards leniency on the part of the staff persons who provided ratings. Median item standard deviation was .75 (range = 1 to 3). Coefficient Alpha internal consistency reliability was $r_\alpha = .95$ for the 23 items, with an average item intercorrelation of $r_{ij} = .48$. The Supervisor Checklist thus seems to be a psychometrically valid measure of staff perceptions of relative worker performance.

The results of the correlational analyses of TPS scores, Supervisor Checklist scores, and IQ measures (PPVT) were encouraging. Two sets of TPS/Supervisor Checklist validity coefficients (Pearson product-moment correlations) were calculated--those for institution-based participants (N = 68) and those for community-based participants (N = 81). The relationship between institution residents' TPS and Supervisor Checklist scores was r = .60, while that for community-based participants was r =.64.

Validity coefficients between TPS/Supervisor Checklist scores were also computed for each site, since the vocational training staff at each site had been asked to rate each worker in terms of other workers at that site. Even though the number of persons on which any of these correlations is based is drastically reduced, this seemed the only appropriate procedure. The range of correlations between Supervisor Checklist and TPS scores was from r = .97 to r = .34, with only two below .50 and a median value of .78.

The two sites with the lowest correlations between TPS and supervisor's scores were the two sites with highest functioning clients; the non-significant correlations were probably due to the restricted range of TPS scores. The large institutions tended to have much lower correlations than the smaller workshops. This appears to be a good indicator of the potential utility of the TPS. At the smaller centers, where there is often individualized training and supervision, the TPS supplies little new information (range: r = .85 to r = .97, median r = .88). However, at the larger institutions, where most placement decisions will have to be made, TPS scores do not agree as strongly with staff judgements (r = .62 at one institution, r = .46 at another).

Correlations between PPVT, Supervisor Checklist, and TPS scores were also calculated; every PPVT/Supervisor Checklist correlation (median r = .40) was appropriately lower than the corresponding TPS/Supervisor rating correlation (median = .61).

A final method of assessing concurrent validity was by contrasted groups (Anastasi, 1973). One such index is the disparity between scores of institutionalized and community-based examinees. Experiences related to TPS tasks are clearly more readily available in community workshop/activity center settings than in typical institutional prevocational training programs. The revised TPS appears quite sensitive to these experience-related effects: the mean total TPS scores for 81 community residents was 71% correct, while that for 68 instituion residents was 45% (standard deviations were similar, 13.2 and 14.0, respectively). Additionally, as can be seen in Table 1, distributions of scores across the possible range of 0 to 60 were distinctly different for the two groups in predicted direction, suggesting that extreme scores of a few do not account for the mean differences.

Summary and Discussion

The revised Trainee Performance Sample will be a psychometrically valid (APA, 1974) instrument to aid in the relative prediction of training time for moderately, severely and profoundly retarded clients. The test attempts to integrate psychometric theory with some of the innovative work in behavioral assessment, criterion-referenced tests, and learning potential assessment. The revised TPS demonstrates more than adequate internal consistency, temporal stability and concurrent validity. TPS scores correlate moderately well with both raw scores on a standardized IQ test and supervisor judgements of worker performance. As predicted, the distribution of TPS scores is higher at community-based workshops than in large institutions. TPS scores correlate more strongly with staff judgements at the smaller workshops than at large institutions, indicating some potential utility for the instrument in deinstitutionalization decisions.

Current activities include (1) criterion-related validity study; (2) factor analysis of test items to revise subscales; and (3) exploration of possible aptitude-treatment interactions.

Criterion-Related Validity

The criterion problem--how ultimately to measure "success"

TABLE 1

Percents of Community/Institution-Based T P S Examinees in Different Score Intervals

Score Intervals	Community-Based [a]	Institution-Based [b]
0-15	4.9	25.0
16-30	13.6	29.4
31-45	29.6	35.3
46-60	51.9	10.3

[a] N = 81
[b] N = 68

in a specific occupation--has been a stumbling block in vocational assessment for normal adults (Wiggins, 1973) and clients of rehabilitation services (Parker & Hansen, 1976). The conventional criterion measure in rehabilitation services--"successful" job placement vs. dropping a client from the caseload--is certainly not an option in services for severely retarded individuals.

Bellamy and Snyder (1976) attempted a more microscopic, precise measure of criterion behavior for the original TPS. A small group of institutionalized clients were trained to perform several vocational tasks using forward or reverse chaining formats. Significant negative correlations (median r = -.72) were found between TPS scores and training trials to criterion, indicating potential predictive validity for the TPS. Though some may argue that this is an overly narrow view of criterion performance, the continual failure of test developers to reliably measure broad range vocational competence (Wiggins, 1973) would lend support to this narrow band methodology. In addition, this particular type of criterion performance is currently the goal of most contemporary vocational training programs for severely retarded adults.

Criterion-related validity of the TPS will be carried out in early 1980. Approximately 30 clients will be trained on 3 prototypical vocational tasks, which systematically involve the task attributes described in Figure 2. Training time will be correlated with the participants' (1) TPS score; (2) PPVT score; and (3) Supervisor's Rating Scale score. The first correlation coefficient will be used as an estimate of criterion-related validity. Fluctuations across tasks will be noted.

A variant of the Campbell-Fiske multi-trait multi-method (cited in Anastasi, 1973) approach will be used to explore the convergent and divergent validity coefficients for the TPS when contrasted with IQ scores and supervisor judgements. It is expected that the TPS will have stronger negative correlations with training time than either a verbal IQ test such as the Peabody, or the judgements of staff. Furthermore, the difference should be larger at institutions than in smaller, more carefully supervised, community workshops.

This approach will offer a relatively rigorous measure of the criterion-related validity, since training conditions will be standardized for all subjects. Interestingly enough, the proposed criterion measure is in many ways similar to the work sample approach advocated by Kulman et al. (1975) except that instruction is incorporated into the work sample. Any interpretations of criterion-related validity must be limited, though, by the nature of the specific training procedures employed, and the type of vocational tasks (small parts assembly) used in the validation study. If these conditions are explicitly stated, reasonably accurate predictions about training time can be made.

Factor Analysis

Two more theoretical concerns seem of importance, and will be explored in the upcoming year. The first of these can be addressed via factor analysis. While administering the test to the 149 participants, we developed a sense that the TPS tapped in a fairly systematic fashion some basic cognitive competencies--such as the ability to match colors, to imitate gestures, to follow simple verbal instructions--that are crucial not only to vocational training, but to general social and intellectual functioning as well. Because the items were selected so as to be appropriate for severely retarded persons, the TPS can tap those abilities far more successfully than can standardized intelligence tests. Current item scores are being factor analyzed to determine which groups of items covary together, and whether there may be logical or instructional relevance to the groupings. Via factor analysis results, subscales will be revised if appropriate, and a determination will be made whether TPS subscale scores can offer training staff useful information on which procedures would be more efficacious for a particular individual. In addition, correlations between factor (or subscale) scores and criterion performance on the three prototypical tasks will be examined.

Aptitude-Treatment Interaction Issues

A final concern is a variant of the aptitude-treatment interaction model. The goal will be to explore whether certain clients learn more efficiently with verbal instructions, whereas others can be trained more easily with use of primarily nonverbal procedures (models and prompts). Because, in our work, both personological variables and criterion behaviors are more behaviorally-defined than in comparable studies in learning/reading disabilities (Salvia & Ysseldyke, 1978) or previous vocational assessment efforts (cf. Neff, 1970; Halpern, 1977), one can at least speculate that aptitude-treatment interactions between clients and training strategies may arise, and certainly should be investigated. As currently conceived, the TPS comes close to measuring what Goldfried and Kent (1972) label "intervening variables," personological variables whose utility is demonstrated with reference to precise, operationally defined situations. The behavioral emphasis in defining both the domain of TPS test items (content validity empirically validated by factor analysis) and criterion measures seems to fit Goldfried and Kent's (1972) criteria.

In summary, we can only speculate as to what will happen to TPS scores in applied settings. If properly validated, a TPS score can offer some information on how quickly a potential vocational client may be trained relative to other similar clients (in a specified domain of tasks). But this number like any other psychological score is subject to error due to examiner and examinee behavior, measurement error and pure chance. There is a concern that this imprecise number will be used to exclude individuals from training as IQ scores were previously used to exclude children from education or that the number will be used to imply individual potential rather than required service efforts. Psychometric theory alone can offer nothing to meet that challenge.

References

Anastasi, A. Psychological testing (4th ed.). New York: Macmillan, 1976.

American Psychological Association. Standards for educational psychological tests. Washington: A.P.A., 1974.

Becker, W., & Engelmann, S. Teaching 3: Evaluation of instruction. Chicago: SRA, 1976.

Bellamy, G. T., Horner, R. H., & Inman, D. P. Vocational habilitation of severely retarded adults: A direct service technology. Baltimore: University Park Press, 1979.

Bellamy, G. T., & Irvin, L. K. Habilitation of the severely and profoundly retarded: An applied research perspective. In C. Cleland, J. Swartz, and L. Talkington (Eds.), The profoundly mentally retarded, Second Annual Conference Proceedings, June, 1976.

Bellamy, T., & Irvin, L. Utility of an instrument for assessing labels used to describe retarded adults. In T. Bellamy, R. Horner, and D. Inman (Eds.), Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program (Vol. II). Eugene, Oregon: University of Oregon Center on Human Development, 1977.

Bellamy, T., & Snyder, S. The trainee performance sample: Toward the prediction of habilitation costs for severely handicapped adults. AAESPH Review, 1976, 1, 17-36.

Budoff, M., & Hamilton, J. Optimizing test performance of moderately and severely mentally retarded adolescents and adults. American Journal of Mental Deficiency, 1976, 81, 49-57.

Buros, O. K. Mental measurements yearbook. Highland Park, N. J.: Gryphon Press, 1972.

Cromwell, R. Ethics, umbrage and the A B C Ds. Minnesota Education, 1976, 2, 42-47.

Dunn, L. Peabody Picture Vocabulary Test Manual. Minneapolis: American Guidance Service, 1973.

Elkin, L. Predicting performance of trainable retardates on experimental workshop tasks. American Journal of Mental Deficiency, 1967, 71, 576-580.

Feuerstein, R. The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments and techniques. In Studies in cognitive modifiability. Jerusalem, 1972.

Gold, M. Research on the vocational habilitation of the retarded: The present, the future. In N. Ellis (Ed.), International review of research in mental retardation (Vol. 6). New York: Academic Press, 1973, 97-148.

Gold, M. Train, don't test. Unpublished manuscript, University of Illinois, 1974.

Goldfried, M. R., & Kent, R. N. Traditional versus behavioral personality assessment: A comparison of methodological and theoretical assumptions. Psychological Bulletin, 1972, 77, 409-420.

Greenwood, C. R., Walker, H. M., & Hops, H. Issues in social interaction/withdrawl assessment. Exceptional Children, 1977, 43, 490-499.

Halpern, A. Principles and practices of measurement in career education for the handicapped. Career Development for Handicapped Individuals, 1977, 1, 13-24.

Haywood, H. C., Filler, J.W., Jr., Shifman, M., & Chatlanat, G. Behavioral assessment in mental retardation. In P. McReynolds (Ed.), Advances in psychological assessment (Vol. 3). Washington, D. C.: Jossey-Bass, 1975, 96-136.

Hively, W. Introduction. In W. Hively and M. Reynolds (Eds.), Domain-referenced testing in special education. Reston, Va.: Council on Exceptional Children, 1975, 1-14.

Irvin, L. K. The general utility of easy-to-hard discrimination training procedures with the severely retarded. Education and Training of the Mentally Retarded, 1976, 11(3), 247-250.

Irvin, L., & Bellamy, T. Long-term retention of vocational skill training by severely retarded individuals. In T. Bellamy, R. Horner, and D. Inman (Eds.), Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program (Vol. II). Eugene, Oregon: University of Oregon Center on Human Development, 1977.

Irvin, L. K., & Halpern, A. The process model of diagnostic assessment. In T. Bellamy, G., O'Connor and O. Karan (Eds.), Vocational rehabilitation of severely handicapped persons: Contemporary service strategies. Baltimore: University Park Press, 1979.

Kulman, H., Stephens, D., Armstrong, C., May, J., Rosenberg, B., Sax, A., Spoonster, J., & Thurgood, K. The tools of vocational evaluation. Vocational Evaluation and Work Adjustment Bulletin, 1975, 8, 49-64.

Levine, S., & Elzey, F. San Francisco Vocational Competency Scale. New York: Pscyhological Corporation, 1968.

Mercer, J. Sociocultural factors in educational labeling. In M. Begab and S. Richardson (Eds.), The mentally retarded: A social science perspective. Baltimore: University Park Press, 1975.

Nunnally, J. Pschometric theory. New York: McGraw-Hill, 1967.

Palmer, R. Pridiction of work performance and work adjustment in mentally retarded adults. Sweden: Scandinavian University Books, 1974.

Parker, R. M., & Hansen, C. E. Aptitude and achievement tests. In B. Bolton (Ed.), Handbook of measurement and evaluation in rehabilitation. Baltimore: University Park Press, 1976.

Payne, J. L. Principles of social science measurement. College Station, Texas: Lytton, 1975.

Popham, W. J. Educational evaluation. Englewood Cliffs, N. J.: Prentice-Hall, 1974.

Rankin, R. Personal communication, University of Oregon, 1978.

Rosenberg, B. The job sample in vocational evaluation (Final Report, Project RD-561). New York: Institute for the Crippled and Disabled, 1967.

Salvia, J., & Ysseldyke, J. E. Assessment in special and remedial education. Boston: Houghton Mifflin, 1978.

Testing orientation and work evaluation in rehabilitation. New York: ICD Rehabilitation and Research Center, 1974.

Throne, J. M. Raw versus standardized intelligence test scores as baselines for assessing effectiveness of treatment: Implications for the mentally retarded. British Journal of Subnormality, 1972, 18, 36-43.

Wiggins, J. S. Personality and prediction: Principles of personality assessment. Reading, Mass.: Addison-Wesley, 1973.

Multiple Assessment Evaluation of Programs for Severely Retarded Adults

Larry K. Irvin, Fred A. Crowell and G. Thomas Bellamy

University of Oregon

This chapter originally appeared in Mental Retardation, 1979, 17(3), 123-128. Reprinted with permission.

Multiple Assessment Evaluation of Programs for Severely Retarded Adults

Recent trends toward deinstitutionalization, community residence, and vocational opportunity for severely retarded persons have created an increased emphasis on appropriate program evaluation models. One such model, the multiple assessment strategy in program evaluation, involves the simultaneous use of several measures to assess outcomes. The purpose of this paper is to describe and examine the application of a multiple assessment strategy within program evaluation of services for retarded adults.

The use of multiple instruments is common in a number of research and clinical contexts (Campbell & Fiske, 1959; Herson & Bellack, 1976), but is seldom encountered in the evaluation of service delivery programs for retarded individuals. The multiple assessment strategy has several advantages over a single instrument approach to program evaluation.

Data quality. The first of these advantages relates to the "believability" of program evaluation data. Some measurement error is inherent in any evaluation instrument, resulting in part from the measurement method and source of data (Cronbach, 1970). The possibility that an identified program outcome reflects this measurement error, rather than actual changes, is reduced when the same outcome is documented through different measurement methods which are susceptible to different error influences (Johnson & Bolstad, 1973). For example, considerably more credence can be given to reported skill improvements of retarded persons if one source of information is corroborated by a second informant, or if such improvements are also indicated by another measure, such as a test or direct behavioral observation.

Multiple outcomes. Perhaps the most obvious reason for using multiple instruments in program evaluation is the need to assess a variety of outcomes in relation to multiple program objectives. It is unlikely that a single measure can be identified which reflects meaningful changes in the quality of normalcy of an individual's lifestyle and environment. At the very least, different instruments are usually needed for assessment of skill development, daily behavior, and characteristics of living environments.

Still another outcome of interest in evaluating service programs is the viability of the program itself, and the extent to which the planned procedures are implemented. Although one instrument might provide a modicum of data on both program characteristics and characteristics of individuals served, it is likely that different instruments would be required for meaningful assessment of these two outcomes.

Multiple users. A neglected but critically important reason for using a multiple assessment strategy in program evaluation is that there are often several potential data users with quite different information needs. For example, the program providing services, the consumers of the services, the agency purchasing the services, the legislative body approving funds for service provision, and the citizenry responsible for electing legislative representatives must all be considered potential users of program evaluation data. Providing data to some of these decision-makers but not to others limits effective participation in program decisions and introduces possible distortions of program benefits. One important difference among users' information needs typically relates to the bandwidth and fidelity characteristics of an instrument (Cronbach, 1970). Program administrators and legislators often need broad bandwidth data which provide a general overview of a wide range of client behaviors, while program planners typically require more detailed, and accurate (i.e., high fidelity) data on a narrower range of behaviors.

An overview. Figure 1 provides a schematic representation of one approach to implementing the multiple evaluation strategy. Listed on the vertical axis are a few outcomes which are typically of interest with respect to services for retarded adults. In selecting a set of measures, it is essential to assess the program outcomes relevant to various decision-makers. Then instrument selection can proceed so that data can be obtained from the client, program, or other source, using tests, ratings and behavior observation instruments as appropriate.

Instrumentation

Two measurement approaches deserve particular attention in relation to multiple assessment evaluation of programs that serve severely retarded individuals: Rating instruments, because they have been used so extensively in program evaluation efforts, (for example, Aanes & Moen, 1976; and Taylor, 1976); and naturalistic observation strategies, because their potential use remains largely untapped. Direct tests of an individual's behavior that measure the extent of response upon demand will not be examined in much detail. The format-and content-appropriateness of almost all currently available tests is limited for evaluation of programs that serve severely retarded individuals. Promising exceptions will be noted.

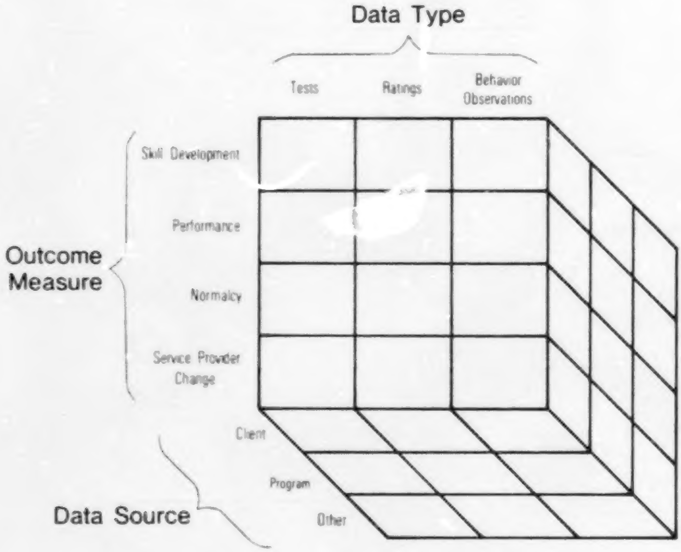


FIGURE 1

Instrumentation Options in Multiple Evaluation

Rating scales. From one perspective, the extent of rating scale usage appears to exceed its utility. From another perspective, though, such instruments are not often used to their fullest advantage; unimaginative use precludes acquisition of important information.

Most users are cognizant of at least some of the problems with rating scales that relate to reliability: acquaintance of rater with ratee, ambiguity of items, halo effect, leniency, error of central tendency, etc. The net effect of any or all of these reliability problems is reduced quality of data, i.e., data that would be different if collected on a different day or by a different person. The wide use of various popular rating scales, despite interobserver reliabilities as low as r = .30, or no reported reliabilities at all, renders the results of evaluation which employs such instrumentation suspect at best and, often, meaningless.

The meaning and validity of item scaling and scoring are other concerns that surface with regard to rating scale usage for purposes of program evaluation. Though items are typically arranged and scored in a hierarchy of skill levels, data are not often reported that support the implicit assumption that successful performance of an item component at a "higher level" is actually dependent on and/or "worth more" than performance at a "lower level" item component. A wide variety of user evaluation needs exist, such as information to guide grouping for training, that cannot be served by ratings because the instruments are not scaled in a manner which facilitates identification of functional differences between individuals.

Another concern involves the degree to which behaviors assessed by rating scales are representative of whole domains of behavior. Groups of items or item components are usually labeled with generic headings such as "Domestic Living," Vocational," etc. The item contents, however, do not often fully justify the headings. A major contributing factor appears to be the limited range of behavior assessed within domains or subdomains of commonly used rating scales. For example, in the Economic Activity, Domestic Activity, and Vocational Activity domains of the AAMD Adaptive Behavior Scale (Nihira, et al., 1974) the numbers of items are only 4, 6, and 3, respectively. Adequate measurement of behavioral repertoires for many program evaluation purposes in these areas often requires more extensive scale content. Again, the consequences for users are that only a limited number of outcomes can be evaluated. On-going formative evaluation, in which a variety of program decisions are continuously based on outcomes of ongoing assessment, cannot be accomplished using such scales.

A final problem that arises when rating scales are used for evaluation of programs serving retarded individuals relates to the primarily descriptive nature of most rating scale instruments. When assessment is accomplished using most rating scales, the resultant measures are commonly limited to static descriptions of the extent of examinee skill, attitude, or trait, as measured by those various, perhaps unreliable and invalid, rating scales. Because severely retarded individuals exhibit low skill or other undesirable behavioral attributes, a far more flexibly useful result would be specific prescriptions for examinee training.

The descriptive nature of most rating scales limits their utility in two important ways. First, specific training outcomes are not easily discernible if broad-based descriptive ratings are employed as measures. The second limit can best be described with two related questions: (a) After measurement has occurred, what has been measured?; (b) So what? Unless such questions can be answered with definite behavioral statements that identify variables over which data users have control, the utility of such instruments for meaningful program evaluation is, at best, limited; only a select group of users can employ the outcome measures meaningfully, i.e., administrative, not training, decision-makers.

Rating scales certainly can be used more imaginatively and to fuller advantage than is currently the typical case. A few available scales demonstrate notable attempts to facilitate accurate prescriptive assessment (Walls, et al., 1976). The Nebraska Client Progress System (1974), and the Behavioral Characteristics Progression (1974) are examples of rating systems that provide prescriptive measures. In addition, Berdine et al. (1977) describe a criterion referenced adaptation of the AAMD Adaptive Behavior Scale that appears to provide functional evaluation data, and Reavis, et al. (1976) describe a useful approach to assessment of severely and profoundly retarded individuals.

Another dimension that rating scale instruments might usefully focus on is process of learning or performing. The typical rating scale results in measures of prior products of a person's learning, i.e., those skills a person brings to an assessment situation. A process focus, on the other hand, is particularly appropriate for assessing persons of low skill level because it results in measures of an individual's mode(s) of acquiring new skills. Valuable estimate of training resources needed could be provided for a given group of retarded individuals.

While no such rating scales are currently available, several relevant test development efforts have occurred recently which could serve well as information sources for guiding conceptualization and development of analogous rating scale development, (see Irvin and Halpern's (1978) recent review).

A final use that could be made of rating scales involves evaluation of program outcomes like "normalization" or "quality of life". Bellamy and Irvin (1977) describe one such application that involves the use of an adjective checklist to measure the extent to which skill acquisition by retarded persons is reflected in the labels which important others use to describe those retarded individuals.

Naturalistic observation instruments (NOIs). In contrast to the widespread application of rating instruments, naturalistic observation techniques have been used very infrequently in program evaluation. A clear discrepancy now exists between the availability of an observation technology and its limited application in the evaluation of community services for retarded adults.

A survey of NOIs used in various research settings yields three major types: continuous recording codes, probe codes, and interval codes. Table 1 presents some examples of these three

TABLE 1

Naturalistic Observation Instruments

Example	Outcome Measures	Instrument Use	Instrument Characteristics
[illegible] (STP Program)	- Appropriate engagement with objects - Positive social - Behavioral complexity	Program effectiveness decisions	- Fixed interval type - Narrow-to-broad bandwidth - 10-second time sample
[illegible] (Landesman-Dwyer)	- Activities - Communication - Interactors - Location	Descriptive/demographic,	- Broad bandwidth - Instrumentation & multiple observers required - Continuous event type - High cost
Activity Code (Bjaanes & Butler)	- Activities	Facility comparisons	- Probe interval type - Low frequency behaviors not recorded - Narrow bandwidth
[illegible] (Taylor & Close)	- Positive behavior - Negative behavior - Normative behavior	Facility comparisons	- Fixed interval type - Narrow bandwidth - 6-second time sample
Social Interaction Code (Berkson & Romer)	- Social behaviors	Descriptive (workshops)	- Continuous type - Broad bandwidth
RAM (Cataldo & Risley)	- Stimulation - Interaction - Activities	Ward design Decisions	- Probe interval type - Low cost - Control of antecedent events - Narrow bandwidth

types and illustrates important characteristics of each. The primary uses of these instruments to date have been in the determination of program effectiveness (Crowell, et al., 1977), and in the comparison of residential facilities (Landesman-Dwyer, Stein,& Sackett, 1976; Close, 1977; Bjaanes & Butler, 1974). Other uses have included program decisions in the organization of institution wards (Cataldo & Risley, 1974), and description of general behaviors (Berkson & Romer, 1977).

In comparison with other available instruments, NOIs have three particularly important features with respect to program evaluation applications. First, they can generally be classified as high fidelity, narrow bandwidth instruments. As such, the focus of measurement is on a small range of outcomes. These outcomes, however, can be indexed with more precision than is possible with broader bandwidth instruments like rating scales. Second, use of NOIs enables collection of data which is not dependent on the verbal skills of the individual being assessed; the largely non-verbal behavior of severely retarded individuals can be assessed to provide direct information on program outcomes. The third advantage of NOIs in program evaluation contexts results from their potential use in measuring both behavior and concurrent environmental events. Reliable monitoring of co-variation over time between specific behavior and stimuli can provide an empirical basis for many treatment decisions in program operation.

Some guidelines for making decisions or choices among NOIs have been reported (Herbert & Attridge, 1975). Suggestions which follow emphasize those criteria which could facilitate use of naturalistic observation in a multiple evaluation context.

One of the major constraints in selecting any instrument is that of cost. While NOIs are usually more expensive than alternative measures, the increased fidelity may offset the cost considerations in many evaluation contexts. NOIs may also be particularly appropriate when coordination among several data users makes it possible to distribute the costs among several agencies or organizations which can benefit from the data obtained. Of course, the ability of multiple users to benefit from one observational data set is determined partially by the coding complexity which results from an instrument's theoretical basis. For maximum cost sharing, instruments are needed in which the complexity of each user's information needs is matched by corresponding aspects of the coding system.

The usual criteria of instrument reliability and validity also apply to the design and selection of NOIs within a multiple evaluation context. As these instruments are used more frequently, issues of data quality will no doubt assume greater importance. Careful attention to issues of scoring, analysis and interpretation of observation data may assist in development of standards of data quality that could facilitate instrument selection (Jones, 1973).

An extensive body of literature awaits the potential user of a naturalistic observation code in the program evaluation context. Numerous methodological articles are available to guide selection and use or new development (for example, Kent & Foster, 1977; and Sackett, 1977).

Use of the Multiple Evaluation Strategy

The following example can serve to illustrate the multiple evaluation strategy. In order to optimize comprehensiveness, quality and utility of data, an evaluator of services to retarded persons might choose:

(a) the AAMD Adaptive Behavior Scales (Nihira, et al., 1974) or another broad bandwidth measure of skill acquisition and performance that involves a rating by program staff;
(b) a naturalistic observation instrument which provides a high fidelity measure of behavioral performance (i.e., behavior observation by some other person not associated with the program;
(c) the Program Analysis of Service Systems (Wolfensberger & Glenn, 1973), to assess normalcy of the environment (a rating by a trained "other"); and
(d) a rating of clients' or their families' satisfaction with services.

With this assessment approach, an evaluator should be able to provide broad bandwidth information to agencies and legislative groups, to check the validity of that information against concurrent measures which involve other data types and sources, and provide high fidelity information to program staff for individual treatment decisions.

As a representation of possible measurement approaches, Figure 1 should also illustrate the need for continued instrument development. At present, reliable instruments are available for measuring each outcome, but these rely heavily on ratings by program staff members or other individuals. Development of alternative measurement methods for each of these outcomes should increase the flexibility and accuracy of evaluation efforts.

The future use of a multiple assessment strategy will no doubt be a function of: perceived and actual benefits in relation to increased costs; the degree to which costs as well as benefits can be shared among a community of potential data users; the availability of design and selection criteria for choosing among instruments; and finally, the degree to which service providers are able to conceptualize and measure such complex outcomes as "quality of life" of service recipients. Current evaluation efforts seldom meet the needs of all important data users, and often employ less than optimal variety of data sources necessary for acquiring evaluative data of high quality and adequate scope. The multiple assessment strategy advocated here may provide one means of responding to the expanding need for program evaluation data.

References

Aanes, D., & Moen, M. Adaptive behavior changes of group home residents. *Mental Retardation*, 1976, *14*(4), 36-40.

Behavioral Characteristics Progression. Palo Alto, California: VORT Corp., P. O. Box 11132, 1974.

Bellamy, T. & Irvin, L. Utility of an instrument for assessing labels used to describe retarded adults. In G. Bellamy, R. Horner, and D. Inman (Eds.), *Habilitation of severely and profoundly retarded adults: Reports from the Specialized Training Program*. Eugene, Oregon: University of Oregon Center on Human Development, 1977, 137-142.

Berdine, W. H., Murphy, M. & Roller, J. D. A criterion-referenced training program based on the ABS: The Oakwood Resident Scale for Training and Evaluating Programs. *Mental Retardation*, 1977, *15*(6), 19-22.

Berkson, G. & Romer, D. *Social behavior of mentally retarded adults*. Paper presented at the annual meeting of the American Association on Mental Deficiency, New Orleans, June 1977.

Bjaanes, A. T. and Butler, E. W. Environmental variation in community care facilities for mentally retarded persons. *American Journal on Mental Deficiency*, 1974, *78*(4), 429-439.

Campbell, D. T. & Fiske, D. W. Convergent and discriminate validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Cataldo, M. T. & Risley, T. R. Evaluation of living environments: he MANIFEST description of ward activities. In P. O. Davison, F. Clark and G. Hamerlynck (Eds.), *Evaluation of Behavioral Programs*. Champaign: Research Press, 1974.

Close, D. Community living for severely and profoundly retarded adults: A group home study. *Education and Training of the Mentally Retarded*, 1977, *12*, 256-262.

Cronbach, L. J. *Essentials of psychological testing*. New York: Harper & Row, 1970.

Crowell, F., Lamson, D., Weissman, N. & Bellamy, T. *NOLE: A naturalistic observation study of living environments of severely retarded vocational trainees. Unpublished manuscript, Center on Human Development,* University of Oregon, Eugene, 1977.

Herbert, J. & Attridge, C. A guide for developers and users of observation systems and manuals. *American Educational Research Journal*, 1975, *12*, 1-20.

Hersen, M. & Bellack, A. *Behavioral Assessment: A Practical Handbook*. London: Pergammon, 1976.

Irvin, L. K. & Halpern, A. S. *The process model of diagnostic assessment*. Paper presented at Timberline Conference on Vocational Habilitation of the Severely Handicapped, University of Oregon, February 1978.

Johnson, S. & Bolstad, O. Methodological issues in naturalistic observation: Some problems and solutions for field research. In L. Hamerlynck, L. Handy and E. Mash (Eds.), *Behavior change: Methodology, concepts and practice*. Champaign: Research Press, 1973.

Jones, R. R. Behavioral observation and frequency data: Problems in scoring, analysis and interpretation. In L. A. Hamerlynck, L. C. Handy and E. J. Mash (Eds.), *Behavior change: Methodology, concepts and practice*. Champaign:

Research press, 1973.

Kent, R. N. & Foster, S. L. Direct observational procedures: Methodological issues in naturalistic settings. In A. Ciminero, K. Calhoun and H. Adams (Eds.), *Handbook of behavioral assessment*. New York: Wiley, 1977.

Landesman-Dwyer, S., Stein, J. & Sackett, G. Group homes for the mentally retarded: An ecological and behavioral study. *Report #2 of Developmental Disabilities Client Placement Study*. Olympia: Department of Social and Health Services, 1976.

Nebraska Client Progress System. Lincoln, Nebraska: Community Education and Development, Inc. 2546 South 48th Street, 1974.

Nihira, K., Foster, R., Shellhaas, M., & Leland, H. *AAMD Adaptive Behavior Scale*. Washington, D.C.: American Association on Mental Deficiency, 1974.

O'Connor, G. *Home is a good place: A national perspective on community residential facilities for the developmentally disabled*. American Association of Mental Deficiency Monograph Series, 1976.

Reavis, H. K., Morrey, J. G. & Hamel, K. Behavioral assessment procedures for community reintegration. *AAESPH Review*, 1976, *1*(4), 37-46.

Sackett, G. P. Measurement in observational research. In G. P. Sackett (Ed.) *Observing behavior, Volume 2: Data collection and analysis methods*. Baltimore: University Park Press, 1977.

Taylor, F. A comparison of adaptive behavior of retarded individuals successfully and unsuccessfully placed in group living homes. *Education and Training of the Mentally Retarded*, February, 1976, 56-64.

Walls, R., Werner, T. & Bacon, A. *Behavior Checklists*. Morgantown, West Virginia: West Virginia University, Rehabilitation Research and Training Center, 1976.

Wolfensberger, W. & Glenn, L. *Pass: A method for the quantitative evaluation of human services*. Toronto, Canada: National Institute on Mental Retardation, 1973.